

RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.*;
 RL Mol. Microbiol. 21:77-96(1996).
 CC -1- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
 DR EMBL: AL031155; CAA20070.1; -.
 DR InterPro: IPR000073; -.
 DR InterPro: IPR000379; -.
 DR Pfam: PF00561; abhydrolase; 1.
 DR Hydrolase.
 KW Hydrolase.
 SQ SEQUENCE 449 AA; 47531 MW; EC74CB1BF3203927 CRC64;

Query Match 81.6%; Score 40; DB 2; Length 449;
 Best Local Similarity 55.6%; Pred. No. 14;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Y 1 ISRPANG 9
 361 LTRYAMGEI 369

RESULT 5
 97

Q9RSB7 PRELIMINARY; PRT; 126 AA.

Q9RSB7; 01-MAY-2000 (TREMBLrel. 13, Created)

01-MAY-2000 (TREMBLrel. 13, Last sequence update)

01-MAR-2001 (TREMBLrel. 16, Last annotation update)

LACTOYLGLUTATHIONE LYASE, PUTATIVE.

DR2208.

Deinococcus radiodurans.

Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.

NCBI_TaxID=1299;

SEQUENCE FROM N.A.

STRAIN-RI;

MEDLINE-20036896; PubMed-10567266;

White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

Yamatheva J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

Kechum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

Fraser C.M.;

*Genome sequence of the radioresistant bacterium Deinococcus

radiodurans RL.*;

Science 286:1571-1577(1999).

EMBL: AE002054; AAF11756.1; -.

TIGR: DR2208; -.

InterPro: IPR000325; -.

Pfam: PF00903; Glyoxalase; 1.

SEQUENCE 126 AA; 13727 MW; 39623D6AB662A18 CRC64;

Query Match 77.6%; Score 38; DB 2; Length 126;

Best Local Similarity 62.5%; Pred. No. 8.2;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Y 1 ISRPANG 8

94 VTRYAMGE 101

RESULT 6

P94767 PRELIMINARY; PRT; 691 AA.

P94767; 01-MAY-1997 (TREMBLrel. 03, Created)

01-MAY-1997 (TREMBLrel. 03, Last sequence update)

01-MAR-2001 (TREMBLrel. 16, Last annotation update)

HGCC PRECURSOR.

HRCC.

Erwinia chrysanthemi.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Pectobacterium.
 OK NCBI_TaxID=556;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-EC16;
 RA Kim J.F., Bauer D.W., Collmer A., Beer S.V.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L39897; AAC31975.1; -.
 DR InterPro: IPR00016; -.
 DR InterPro: IPR003522; -.
 DR Pfam: PF00263; Bcc_GSPproteins; 1.
 DR PRINTS: PR01337; TYPE3OMGPOT.
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 1 31
 74829 MW; 25EC4676959EB2F CRC64;

Query Match 77.6%; Score 38; DB 2; Length 691;
 Best Local Similarity 75.0%; Pred. No. 51;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 SRFAMGE 9
 142 SRFAMGE 149

RESULT 7
 P73526

P73526 PRELIMINARY; PRT; 289 AA.

P73526; 01-FEB-1997 (TREMBLrel. 02, Created)

01-FEB-1997 (TREMBLrel. 02, Last sequence update)

01-JUN-2000 (TREMBLrel. 14, Last annotation update)

HYPOTHEICAL 32.6 KDA PROTEIN.

GN SL1348.

Synechocystis sp. (strain PCC 6803);

Bacteria; Cyanobacteria; Chroococcales; Synechocystis;

NCBI_TaxID=1148;

SEQUENCE FROM N.A.

MEDLINE-97061201; PubMed-8905231;

Kaneko T., Sato S., Kocant H., Tanaka A., Asamizu E., Nakamura Y.,

Miyajima N., Hirosewa M., Sugitani M., Sasamoto S., Kimura T.,

Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,

Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,

Tabata S.;

*Sequence analysis of the genome of the unicellular cyanobacterium

Synechocystis sp. strain PCC6803. II. Sequence determination of the

entire genome and assignment of potential protein-coding regions.*;

DNA Res. 3:109-136(1996).

EMBL: D90907; BAA17566.1; -.

InterPro: IPR001450; -.

Pfam: PF00037; fer4; 1.

PROSITE: PS00196; 4FEAS_FERREDOXIN; UNKNOWN_1.

HYPOTHETICAL protein.

SEQUENCE 289 AA; 32551 MW; DE79651E9AB31DA9 CRC64;

Query Match 75.5%; Score 37; DB 2; Length 289;

Best Local Similarity 85.7%; Pred. No. 31;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRPANG 7

73 ISRPANG 79

RESULT 8

O91602 PRELIMINARY; PRT; 394 AA.

O91602; 01-NOV-1996 (TREMBLrel. 01, Created)

01-NOV-1996 (Tremblrel. 01, Last sequence update)
 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 POLY(A) POLYMERASE.
 PAP.
 Xenopus laevis (African clawed frog).
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 (1)
 SEQUENCE FROM N.A.
 TISSUE-OVARY;
 MEDLINE-95166227; PubMed-7862135;
 Gebauer F., Richter J.D.;
 "Cloning and characterization of a Xenopus poly(A) polymerase.";
 Mol. Cell. Biol. 15:1422-1430(1995).
 EMBL; U23456; AAA64708.1;
 InterPro: IPR001201;
 InterPro: IPR002934;
 Pfam: PF01909; NTP_transf. 2; 1.
 SEQUENCE 394 AA; 45326 MW; 74E4684EF3C1B967 CRC64;

Query Match
 Best Local Similarity 75.5%; Score 37; DB 13; Length 394;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 1 ISPRAMEV 9
 :::::
 376 LAMFAMEI 384

SUFT 9
 0205 PRELIMINARY; PRT; 164 AA.
 092J05;
 01-MAY-1999 (Tremblrel. 10, Created)
 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 PHOTON-TRANSLLOCATING ATPASE B SUBUNIT.
 UNCF.
 Streptococcus sanguis.
 Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 Streptococcus.
 NCBI_TaxID=1305;
 (1)
 SEQUENCE FROM N.A.
 STRAIN-10904;
 Kuhnert W.L., Sells J., Quivey R.G. Jr.;
 "Cloning and Structure/Function Studies of a Proton-Translocating
 ATPase from the Oral Streptococci Streptococcus sanguis."
 Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 EMBL; AF001955; AAD00914.1;
 InterPro: IPR002146;
 Pfam: PF00430; ATP_synt. B; 1.
 SEQUENCE 164 AA; 17595 MW; 37ABD5392FE6212D CRC64;

Query Match
 Best Local Similarity 73.5%; Score 36; DB 2; Length 164;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 1 ISPRAMEV 9
 :::::
 26 IKRFAMGNI 34

SUFT 10
 05B4 PRELIMINARY; PRT; 169 AA.
 0905B4;
 01-MAY-2000 (Tremblrel. 13, Created)
 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE F43C11.2 PROTEIN.
 GN F43C11.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peldoderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 (1)
 SEQUENCE FROM N.A.
 STRAIN-BRISTOL N2;
 RX MEDLINE-99069613; PubMed-9851916;
 RA None.
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 Science 282:2012-2018(1998).
 RN 12
 SEQUENCE FROM N.A.
 RP STRAIN-BRISTOL N2;
 RC Gattung S., Du H., Scheet P., Hawrysko C.;
 RA "The sequence of C. elegans cosmid F43C11."
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN 13
 SEQUENCE FROM N.A.
 RP STRAIN-BRISTOL N2;
 RC Waterston R.;
 RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AC017117; AAF16612.1;
 DR InterPro: IPR003326;
 DR Pfam: PF02343; DUF130; 1.
 SEQUENCE 169 AA; 19239 MW; 3E0BF405CFB19505 CRC64;

Query Match
 Best Local Similarity 73.5%; Score 36; DB 5; Length 169;
 Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 0Y 1 ISPRAMEV 9
 :::::
 Db 106 VARYSGOV 114

RESULT 11
 0900E5 PRELIMINARY; PRT; 228 AA.
 AC 0900E5;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE NONSTRUCTURAL PROTEIN 1 (FRAGMENT).
 GN NSI.
 OS Influenza A virus (A/Quail/Arkansas/29209-1/93 (H9N2)).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 OC Influenza virus A and B group; Influenza A viruses; Influenza A virus.
 NCBI_TaxID=97395;
 (1)
 SEQUENCE FROM N.A.
 RP STRAIN-A/QUAIL/ARKANSAS/29209-1/93;
 RX MEDLINE-99362763; PubMed-10430948;
 RA Guan Y., Shortridge K.F., Kraus S., Webster R.G.;
 RT "Molecular characterization of H9N2 influenza viruses: were they the
 donors of the 'internal' genes of H5N1 viruses in Hong Kong?";
 DL Proc. Natl. Acad. Sci. U.S.A. 96:9363-9367(1999).
 DR EMBL; AF156484; AAD52959.1;
 DR HSSP: P03495; IAIL.
 DR InterPro: IPR000256;
 DR Pfam: PF00600; Flu_NSI; 1.
 KW Nonstructural protein.
 FT NON_TER
 SEQUENCE 228 AA; 25806 MW; 363D913707208C08 CRC64;

Query Match
 Best Local Similarity 73.5%; Score 36; DB 14; Length 228;
 Matches 85.7%; Pred. No. 37;

39XAN3
ID 09XAN3 PRELIMINARY: PRT: 375 AA.
1C 01-NOV-1999 (TREMUREL. 12, Created)
2C 01-NOV-1999 (TREMUREL. 12, Last sequence update)
3C 01-MAR-2001 (TREMUREL. 16, Last annotation update)
4C PUTATIVE 3-OXOADIPATE ENOL-LACTONE HYDROLASE/4-CARBOXYMUCONOLACTONE
5C DECARBOXYLASE.
6C SC66.07C.
7C Streptomyces coelicolor.
8C Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
9C Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
10C NCBI_TaxID:1902;
11C [1]
12C SEQUENCE FROM N.A.
13C STRAIN-A3(2);
14C Seeger S., Harris D.;
15C "A set of ordered cosmids and a detailed genetic and physical map for
16C the 8 Mb Streptomyces coelicolor A3(2) chromosome."
17C Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
18C [2]
19C SEQUENCE FROM N.A.
20C STRAIN-A3(2);
21C MEDLINE-97000351; PubMed-8843436;
22C Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
23C Kinashi H., Hopwood D.A.;
24C "A set of ordered cosmids and a detailed genetic and physical map for
25C the 8 Mb Streptomyces coelicolor A3(2) chromosome."
26C M1. Microbiol. 21:77-96(1996).
27C EMBL; AL079355; CAB45571.1;
28C InterPro: IPR000073;
29C InterPro: IPR000379;
30C InterPro: IPR003089;
31C Pfam: PF00361; abhydrolase; 1.
32C PRINTS: PR0011; ABHYDROLASE.
33C Hydrolase.
34C SEQUENCE 375 AA; 39873 MW; 67176A8F3E61EECC CRC64;

Query Match 91.8%; Score 45; DB 2; Length 375;
1C Local Similarity 77.8%; Pred. No. 1.3;
2C Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 ISRPAGEV 9
111:1111;
b 286 ISRPAGEI 294

RESULT 3
9PGJ3 PRELIMINARY: PRT: 250 AA.
0C 09PGJ3;
1C 01-NOV-2000 (TREMUREL. 15, Created)
2C 01-OCT-2000 (TREMUREL. 15, Last sequence update)
3C 01-MAR-2001 (TREMUREL. 16, Last annotation update)
4C NADH-UBIQUINONE OXIDOREDUCTASE, NQO5 SUBUNIT.
5C XE0307.
6C Xylella fastidiosa.
7C Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
8C Xylella.
9C NCBI_TaxID-2371;
10C [1]
11C SEQUENCE FROM N.A.
12C STRAIN-9A5C;
13C MEDLINE-20365717; PubMed-10910347;
14C Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
15C Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.N.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Falcinelli A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Faria J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Landais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madalita A.M.B.N., Madalita H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Meon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
RA Quaglini R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva M.A., Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsunako M.H.,
RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RA "The genome sequence of the plant pathogen Xylella fastidiosa."
RA Nature 406:151-159(2000).
RA EMBL; AE003884; AAF83118.1;
RA InterPro: IPR001268;
RA DR Pfam: PF00329; complex_30kd; 1.
RA PROSITE; PS00542; COMPLEX_30K; 1.
RA SEQUENCE 250 AA; 28081 MW; 2A96C7D0C981F8DB CRC64;

Query Match 81.6%; Score 40; DB 2; Length 250;
1C Local Similarity 75.0%; Pred. No. 7.3;
2C Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Y 1 ISRPAGE 8
111:1111;
Db 90 VGRPAGE 97

RESULT 4
086608 PRELIMINARY: PRT: 449 AA.
ID 086608;
AC 086608;
DT 01-NOV-1998 (TREMUREL. 08, Created)
DT 01-NOV-1998 (TREMUREL. 08, Last sequence update)
DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)
DE 3-OXOADIPATE ENOL-LACTONE HYDROLASE/4-CARBOXYMUCONOLACTONE
DE DECARBOXYLASE.
GN SC3A7.07.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID-1902;
1C [1]
2C SEQUENCE FROM N.A.
3C STRAIN-A3(2);
4C Oliver K., Harris D.;
5C Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
6C [2]
7C SEQUENCE FROM N.A.
8C STRAIN-A3(2);
9C Parkhill J., Barrell B.G., Rajandream M.A.,
10C Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
11C [3]
12C SEQUENCE FROM N.A.
13C STRAIN-A3(2);
14C MEDLINE-97000351; PubMed-8843436;
15C Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
16C Kinashi H., Hopwood D.A.;
17C "A set of ordered cosmids and a detailed genetic and physical map for

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

M protein - protein search, using sw model

Run on: October 18, 2001, 16:53:10 ; Search time 33.31 Seconds
(without alignments)
35.747 Million cell updates/sec

Title: US-09-646-579-2

Effect score: 49
Sequence: 1 ISRFAMGEV 9

Corling table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

num DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SPTREMBL.16:**

1: sp.archaea:**
2: sp.bacteria:**
3: sp.fungi:**
4: sp.human:**
5: sp.invertebrate:**
6: sp.mammal:**
7: sp.mhc:**
8: sp.organelle:**
9: sp.phage:**
10: sp.plant:**
11: sp.podent:**
12: sp.unclassified:**
13: sp.vertebrate:**
14: sp.virus:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	91.8	373	2	Q9XD79
2	45	91.8	375	2	Q9XAN3
3	40	81.6	250	2	Q9PGJ3
4	40	81.6	449	2	Q86608
5	38	77.6	126	2	Q9RSB7
6	38	77.6	691	2	P94767
7	37	75.5	289	2	P73526
8	37	75.5	164	13	Q91602
9	36	73.5	169	2	Q92105
10	36	73.5	169	2	Q9U584
11	36	73.5	228	14	Q9U0E5
12	36	73.5	229	14	Q9Q0E3
13	36	73.5	230	14	Q41651
14	36	73.5	230	14	Q41653
15	36	73.5	230	14	Q41654
16	36	73.5	230	14	Q41655
17	36	73.5	230	14	Q41659
18	36	73.5	230	14	Q41661
19	36	73.5	230	14	Q41667

20	36	73.5	230	14	Q42067	042067 influenza a
21	36	73.5	230	14	Q42082	042082 influenza a
22	36	73.5	230	14	Q04262	Q04262 influenza a
23	36	73.5	230	14	Q09702	Q09702 influenza a
24	36	73.5	230	14	Q09704	Q09704 influenza a
25	36	73.5	230	14	Q82804	Q82804 influenza a
26	36	73.5	230	14	Q92649	Q92649 influenza a
27	36	73.5	230	14	Q92650	Q92650 influenza a
28	36	73.5	230	14	Q92785	Q92785 influenza a
29	36	73.5	230	14	Q82813	Q82813 influenza a
30	36	73.5	230	14	Q89283	Q89283 influenza a
31	36	73.5	230	14	Q9WC20	Q9WC20 influenza a
32	36	73.5	230	14	Q9WC19	Q9WC19 influenza a
33	36	73.5	230	14	Q9WC18	Q9WC18 influenza a
34	36	73.5	230	14	Q9WC16	Q9WC16 influenza a
35	36	73.5	230	14	Q9WC15	Q9WC15 influenza a
36	36	73.5	230	14	Q9WC14	Q9WC14 influenza a
37	36	73.5	230	14	Q9WC13	Q9WC13 influenza a
38	36	73.5	230	14	Q9WC12	Q9WC12 influenza a
39	36	73.5	230	14	Q9WC10	Q9WC10 influenza a
40	36	73.5	230	14	Q9Q0L6	Q9Q0L6 influenza a
41	36	73.5	230	14	Q9EA72	Q9EA72 influenza a
42	36	73.5	230	14	Q9EA66	Q9EA66 influenza a
43	36	73.5	230	14	Q9DG10	Q9DG10 influenza a
44	36	73.5	272	2	Q54375	Q54375 lactococcus
45	36	73.5	451	5	Q9TXV7	Q9TXV7 caenorhabdi

ALIGNMENTS

RESULT 1
ID Q9XD79 PRELIMINARY; PRT: 373 AA.
AC Q9XD79;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE 4-CARBOXYMOCOLACTONE DECARBOXYLASE/3-OXOADIPATE ENOL-LACTONE
DE HYDROLASE.
GN PCAL.
OS Streptomyces sp. 2065.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=86383;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2065;
RA Yang K., Iwagami S., Davies J.E.;
RT "A. protocatechnate catabolic gene cluster cloned from Streptomyces sp.
RT 2065.";
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
DR EMBL: AF109386; AAD0815.1;
DR InterPro: IPR000073;
DR InterPro: IPR00379;
DR Pfam: PF00561; abhydrolase; 1.
DR PRINTS: PR00111; ABHYDROLASE.
KW Hydrolase.
SQ SEQUENCE 373 AA: 39563 MW; F11D3017D7A524DC CRC64;

Query Match 91.8%; Score 45; DB 2; Length 373;
Best Local Similarity 77.8%; Pred. No. 1.3;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMGEV 9
|||:||||:
Db 288 ISRYAMGEI 296

RESULT 2

AC P06906; 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MYELIN BASIC PROTEIN (MBP).
 GN MBP.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.
 OX NCBI_TaxId=9598;
 RN [1]
 RP PRELIMINARY SEQUENCE.
 RX MEDLINE=76009821; PubMed=51459;
 RA Westall F.C., Thompson M., Katter S.S.;
 RT "The proposed sequence of the encephalitogenic protein from
 chimpanzee brain."
 RL Life Sci. 17:219-223(1975).
 CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE
 OF MYELIN.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
 CC -1- SIMILARITY: BELONGS TO THE MYELIN BASIC PROTEIN FAMILY.
 CC PIR: A03139; MCEB.
 DR InterPro; IPR000548;
 DR Pfam: PF01669; Myelin_MBP; 1.
 DR PRINTS: PR00212; MYELINMBP.
 DR PROSITE: PS00569; MYELIN_MBP; 1.
 KW Myelin; Structural protein; Acetylation; Methylation; Phosphorylation;
 KW Autoimmune encephalomyelitis.
 FT MOD_RES 1 ACETYLATION.
 FT MOD_RES 1 METHYLATION (BY SIMILARITY).
 SQ SEQUENCE 171 AA; 18560 MW; EPEFDS9DEG933293 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 171;
 Best Local Similarity 71.4%; Pred. No. 8.7;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ISRPAWG 7
 Db 111 ISRFSWG 117

Search completed: October 18, 2001, 16:52:30
 Job time: 115 sec

FT INIT_MET 0 0
 FT MOD_RES 1 1
 FT CONFLICT 104 104
 FT CONFLICT 46 47
 FT CONFLICT 124 124
 SQ SEQUENCE 127 AA; 14080 MW; B4C9F33C19A9E137 CRC64;

Query Match
 Best Local Similarity 71.4%; Score 35; DB 1; Length 127;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRRFAMG 7
 :|||:|
 Db 108 LSRFSWG 114

RESULT 13

MBP_CAVPO STANDARD; PRT; 167 AA.

AC P25188;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MYELIN BASIC PROTEIN (MBP).
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
 NCBI_TaxID=10141;
 RX MEDLINE=84215086; PubMed=6202840;
 RA Dehler G.E., Martenson R.E., Krutzsch H.C., Kles M.W.;
 RT "Sequence of guinea pig myelin basic protein.";
 RL J. Neurochem. 43:100-105(1984).
 RN [2]
 RP SEQUENCE OF 45-87.
 RA Shapira R., McKeenally S.S., Chou F., Kibler R.F.;
 RT "Encephalitogenic fragment of myelin basic protein. Amino acid
 sequence of bovine, rabbit, guinea pig, monkey, and human fragments.";
 RL J. Biol. Chem. 246:4630-4640(1971).
 CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE
 OF MYELIN.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
 CC -1- SIMILARITY: BELONGS TO THE MYELIN BASIC PROTEIN FAMILY.
 DR PIR: A37246; A37246.
 DR PIR: C92087; C92087.
 L InterPro: IPR000548;
 DR Pfam: PF01669; Myelin_MBP; 1.
 DR PRINTS: PR00212; MYELIN_MBP; 1.
 DR PROSITE: PS00569; MYELIN_MBP; 1.
 KM Myelin; Structural protein; Acetylation; Methylation; Phosphorylation;
 KW Autoimmune encephalomyelitis.
 FT MOD_RES 1 1
 FT MOD_RES 106 106
 FT MOD_RES 106 106
 FT MOD_RES 106 106
 FT DOMAIN 114 122
 FT INDICES EXPERIMENTAL AUTOIMMUNE
 ENCEPHALOMYELITIS.
 SQ SEQUENCE 167 AA; 18213 MW; 866D31F5EACFEA6 CRC64;

Query Match
 Best Local Similarity 71.4%; Score 35; DB 1; Length 167;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRRFAMG 7
 :|||:|
 Db 110 LSRFSWG 116

RESULT 14
 MBP_BOVIN STANDARD; PRT; 169 AA.

AC P02687;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MYELIN BASIC PROTEIN (MBP) (MYELIN AI PROTEIN).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovine; Bos.
 NCBI_TaxID=9913;
 RX MEDLINE=72007306; PubMed=5096093;
 RA Eylar E.H., Brostoff S.W., Hashim G., Caccam J., Burnett P.;
 RT "Basic AI protein of the myelin membrane. The complete amino acid
 sequence.";
 RL J. Biol. Chem. 246:5770-5784(1971).
 RN [2]
 RP REVISION.

RX MEDLINE=74070688; PubMed=4129204;
 RA Brostoff S.W., Reuter M., Hitchens M., Eylar E.H.;
 RT "Specific cleavage of the AI protein from myelin with cathepsin D.";
 RL J. Biol. Chem. 249:559-567(1974).
 RN [3]
 RP SEQUENCE OF 43-87.
 RA Shapira R., McKeenally S.S., Chou F., Kibler R.F.;
 RT "Encephalitogenic fragment of myelin basic protein. Amino acid
 sequence of bovine, rabbit, guinea pig, monkey, and human fragments.";
 RL J. Biol. Chem. 246:4630-4640(1971).
 RN [4]
 RP METHYLATION.
 RX MEDLINE=71153946; PubMed=4994464;
 RA Brostoff S.W., Eylar E.H.;
 RT "Localization of methylated arginine in the AI protein from myelin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 68:765-769(1971).
 RN [5]
 RP SYNTHESIS OF ALLERGIC ENCEPHALOMYELITIS INDUCING REGION.
 RX MEDLINE=70178977; PubMed=5442707;
 RA Eylar E.H., Caccam J., Jackson J.J., Westall F.C., Robinson A.B.;
 RT "Experimental allergic encephalomyelitis: synthesis of
 disease-inducing site of the basic protein.";
 RL Science 168:1220-1223(1970).
 CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE
 OF MYELIN.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
 CC -1- SIMILARITY: BELONGS TO THE MYELIN BASIC PROTEIN FAMILY.
 DR PIR: A03140; MBP.
 DR InterPro: IPR000548;
 DR Pfam: PF01669; Myelin_MBP; 1.
 DR PRINTS: PR00212; MYELIN_MBP; 1.
 DR PROSITE: PS00569; MYELIN_MBP; 1.
 KM Myelin; Structural protein; Acetylation; Methylation; Phosphorylation;
 KW Autoimmune encephalomyelitis.
 FT MOD_RES 1 1
 FT MOD_RES 106 106
 FT MOD_RES 106 106
 FT DOMAIN 114 122
 FT INDICES EXPERIMENTAL AUTOIMMUNE
 ENCEPHALOMYELITIS.
 SQ SEQUENCE 169 AA; 18323 MW; 8E157B7A1978484 CRC64;

Query Match
 Best Local Similarity 71.4%; Score 35; DB 1; Length 169;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRRFAMG 7
 :|||:|
 Db 110 LSRFSWG 116

RESULT 15
 MBP_PANTR STANDARD; PRT; 171 AA.

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M55466; AAA43132.1; -
DR HSSP; P03495; IAIL.
DR InterPro; IPR000256; -
DR Pfam; PF00600; Flu.NS1.1.
KW Nonstructural protein: Alternative splicing.
SQ SEQUENCE 230 AA; 26076 MW; 3FDE273901ED1D7C CRC64;

Query Match 73.5%; Score 36; DB 1; Length 230;
Best Local Similarity 85.7%; Pred.No. 7.5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
DB 198 IORFAMG 204

RESULT 11
RSP4_CHLRE STANDARD; PRT; 465 AA.
AC 001656;

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE FLAGELLAR RADIAL SPOKE PROTEIN 4.

GN RSP4.

CS Chlamydomonas reinhardtii.
CC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
CC Chlamydomonadaceae; Chlamydomonas.

CX NCBI_TaxID=3055;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-21GR;
RX MEDLINE-92375065; PubMed-1508197;

RT Curry A.M., Williams B.D., Rosenbaum J.L.;
*Sequence analysis reveals homology between two proteins of the
RT flagellar radial spoke.*

RL Mol. Cell. Biol. 12:3967-3977(1992).

CC -1- FUNCTION: FLAGELLAR RADIAL SPOKES CONTRIBUTE TO THE REGULATION
CC OF DYNEIN ARM ACTIVITY AND THUS THE PATTERN OF FLAGELLAR BENDING.
CC THEY CONSIST OF A THIN STALK, WHICH IS ATTACHED TO THE A SUBFIBER
CC OF THE OUTER DOUBLET MICROTUBULE, AND A BULBOUS HEAD, WHICH IS
CC ATTACHED TO THE STALK AND APPEARS TO INTERACT WITH THE
CC PROJECTIONS FROM THE CENTRAL PAIR OF MICROTUBULES.

CC SUBUNIT: THE RADIAL SPOKE HEAD IS MADE OF FIVE DIFFERENT
CC POLYPEPTIDES (RSP1, RSP4, RSP6, RSP9, AND RSP10).

CC -1- SUBCELLULAR LOCATION: RADIAL SPOKE.

CC -1- SIMILARITY: TO THE FLAGELLAR RADIAL SPOKE PROTEIN 6.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M67526; AAA33092.1; -
DR PIR; A44498; A44498.

DR FLAGELLA.
SQ SEQUENCE 465 AA; 49798 MW; A23AFB030CDB3E29 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 465;
Best Local Similarity 100.0%; Pred.No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FAMGEV 9
DB 438 FAMGEV 443

RESULT 12
MBP_RAT STANDARD; PRT; 127 AA.
AC P02688;

DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE MYELIN BASIC PROTEIN S (MBP S).

GN MBP.

OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;

CX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE-87026249; PubMed-2429678;

RA Schach M., Budzinski R.M., Stoffel W.;
*Cloned proteolipid protein and myelin basic protein cDNA.
RT Transcription of the two genes during myelination.*

RL Biol. Chem. Hoppe-Seyler 367:825-834(1986).

GN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE-84026484; PubMed-6194889;

RA Roach A., Boylan K., Horvath S., Prusiner S.B., Hood L.E.;
*Characterization of cloned cDNA representing rat myelin basic
RT protein: absence of expression in brain of shiverer mutant mice.*

RL Cell 34:799-806(1983).

RN [3]

RP SEQUENCE.
RX MEDLINE-75127359; PubMed-4141893;

RA Dunkley P.R., Carnegie P.R.;
*Amino acid sequence of the smaller basic protein from rat brain
RT myelin.*

RL Biochem. J. 141:243-255(1974).

RN [4]

RP SEQUENCE OF 45-85.
RX MEDLINE-73180720; PubMed-4122324;

RA McFarlane D.E., Blank S.E., Kibler R.F., McNeally S., Shapiro R.;
*Experimental allergic encephalomyelitis in the rat: response to
RT encephalitogenic proteins and peptides.*

RL Science 179:478-480(1973).

CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE
CC OF MYELIN.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.

CC -1- ALTERNATIVE PRODUCTS: RATS HAVE TWO MYELIN BASIC PROTEINS. THE
CC SMALLER ONE, SHOWN HERE, IS MISSING 40 RESIDUES (FOLLOWING RESIDUE
CC 113 OR 114) WITH RESPECT TO THE LARGER ONES FROM OTHER SPECIES.

CC -1- SIMILARITY: BELONGS TO THE MYELIN BASIC PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M25889; AAA41575.1; -
DR EMBL; K00512; -; NOT_ANNOTATED_CDS.

DR PIR; A03142; MBRTS.
DR PIR; B24351; B24351.

DR PIR; A21062; A21062.

DR InterPro; IPR000548; -
DR Pfam; PF01669; Myelin_MBP.1.

DR PRINTS; PR00212; MYELINMBP.
DR POSITE; PS00569; MYELIN_MBP.1.
KW Myelin; Structural protein; Acetylation; Methylation; Phosphorylation;
KW Autoimmune encephalomyelitis; Alternative splicing.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M55370; AAA35559.1; -
 CC PIR: G32663; MNTA8.
 CC HSSP: P03495; IAIL.
 CC InterPro: IPR000256; -
 CC Pfam: PF00600; Flu_NS1; 1.
 CC Nonstructural protein: Alternative splicing.
 CC SEQUENCE 230 AA; 26071 MW; D738A943C8FFBC84 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 230;
 Best Local Similarity 85.7%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ISRFAMG 7
 1
 198 IORFAMG 204

RESULT 8
 VNS1_IATRB STANDARD; PRT; 230 AA.
 ID VNS1_IATRB
 AC P30910;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE NONSTRUCTURAL PROTEIN NS1.
 OS Influenza A virus (strain A/Turkey/Bethlehem-Gillite/1492-B/82).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 CC Influenza virus A and B group; Influenza A viruses;
 CC Influenza A virus;
 CC NCBI_TaxID=31663;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=91306439; PubMed=1830182;
 CC Ludwig S., Schultz U., Mandler J., Fitch W.M., Scholtissek C.;
 CC "Phylogenetic relationship of the nonstructural (NS) genes of
 CC Influenza A viruses.";
 CC Virology 183:566-577(1991).
 CC -1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF SEGMENT 8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M55467; AAA43124.1; -
 CC HSSP: P03495; IAIL.
 CC InterPro: IPR000256; -
 CC Pfam: PF00600; Flu_NS1; 1.
 CC Nonstructural protein: Alternative splicing.
 CC SEQUENCE 230 AA; 26174 MW; 3FDE27220050701C CRC64;

Query Match 73.5%; Score 36; DB 1; Length 230;
 Best Local Similarity 85.7%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ISRFAMG 7
 1
 198 IORFAMG 204

RESULT 9
 VNS1_IATKC STANDARD; PRT; 230 AA.
 ID VNS1_IATKC
 AC P30911;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE NONSTRUCTURAL PROTEIN NS1.
 OS Influenza A virus (strain A/Turkey/Canada/63).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 CC Influenza virus A and B group; Influenza A viruses;
 CC Influenza A virus;
 CC NCBI_TaxID=31664;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=91306439; PubMed=1830182;
 CC Ludwig S., Schultz U., Mandler J., Fitch W.M., Scholtissek C.;
 CC "Phylogenetic relationship of the nonstructural (NS) genes of
 CC Influenza A viruses.";
 CC Virology 183:566-577(1991).
 CC -1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF SEGMENT 8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M55468; AAA43138.1; -
 CC HSSP: P03495; IAIL.
 CC InterPro: IPR000256; -
 CC Pfam: PF00600; Flu_NS1; 1.
 CC Nonstructural protein: Alternative splicing.
 CC SEQUENCE 230 AA; 25964 MW; 41BBA9C632C1057 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 230;
 Best Local Similarity 85.7%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ISRFAMG 7
 1
 198 IORFAMG 204

RESULT 10
 VNS1_IATRT STANDARD; PRT; 230 AA.
 ID VNS1_IATRT
 AC P30912;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE NONSTRUCTURAL PROTEIN NS1.
 OS Influenza A virus (strain A/Tern/Turkmenia/18/72).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 CC Influenza virus A and B group; Influenza A viruses;
 CC Influenza A virus;
 CC NCBI_TaxID=11477;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=91306439; PubMed=1830182;
 CC Ludwig S., Schultz U., Mandler J., Fitch W.M., Scholtissek C.;
 CC "Phylogenetic relationship of the nonstructural (NS) genes of
 CC Influenza A viruses.";
 CC Virology 183:566-577(1991).
 CC -1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF SEGMENT 8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```

CC -----
DR EMBL: J02105; AAA43509.1;
DR EMBL: M25372; AAA43529.1;
DR PIR: A04092; MNIV16.
DR PIR: C32663; MNIV16.
DR HSSP: P03495; IAIL.
DR InterPro: IPR000256;
DR Pfam: PF00600; Flu_NS1; 1.
KW Nonstructural protein; Alternative splicing.
SQ SEQUENCE 230 AA; 26049 MW; DE875F03C8F08FB8 CRC64;

Query Match
Best Local Similarity 73.5%; Score 36; DB 1; Length 230;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ISRFAMG 7
1 |||||
198 IORFAMG 204

LT 5
VNS1_IAMA6 STANDARD; PRT; 230 AA.
VNS1_IAMA6
01-JAN-1990 (Rel. 13, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
01-MAY-1992 (Rel. 22, Last annotation update)
NONSTRUCTURAL PROTEIN NS1.
Influenza A virus (strain A/Mallard/Alberta/88/76).
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
Influenza virus A and B group; Influenza A viruses;
NCBI_TaxID=11430;
[1]
SEQUENCE FROM N.A.
MEDLINE=89299445; PubMed=2525836;
Treator J.J., Snyder M.H., London W.T., Murphy B.R.;
"The B allele of the NS gene of avian influenza viruses, but not the
A allele, attenuates a human influenza A virus for squirrel
monkeys.";
Virology 171:1-9(1989).
-1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
ALTERNATIVE SPLICING OF SEGMENT 8.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: M25372; AAA43531.1;
PIR: A32663; MNIV16.
HSSP: P03495; IAIL.
InterPro: IPR000256;
Pfam: PF00600; Flu_NS1; 1.
KW Nonstructural protein; Alternative splicing.
SQ SEQUENCE 230 AA; 26021 MW; 39BC9521FAFA9590 CRC64;

Query Match
Best Local Similarity 73.5%; Score 36; DB 1; Length 230;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ISRFAMG 7
1 |||||
198 IORFAMG 204

RESULT 6
VNS1_IAP11 STANDARD; PRT; 230 AA.
VNS1_IAP11
01-JAN-1990 (Rel. 13, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
01-MAY-1992 (Rel. 22, Last annotation update)
NONSTRUCTURAL PROTEIN NS1.
Influenza A virus (strain A/Pintail/Alberta/121/79).
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
Influenza virus A and B group; Influenza A viruses;
NCBI_TaxID=11450;
[1]
SEQUENCE FROM N.A.
MEDLINE=89299445; PubMed=2525836;
Treator J.J., Snyder M.H., London W.T., Murphy B.R.;
"The B allele of the NS gene of avian influenza viruses, but not the
A allele, attenuates a human influenza A virus for squirrel
monkeys.";
Virology 171:1-9(1989).
-1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
ALTERNATIVE SPLICING OF SEGMENT 8.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

ID VNS1_IAP11 STANDARD; PRT; 230 AA.
AC P13141;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE NONSTRUCTURAL PROTEIN NS1.
OS Influenza A virus (strain A/Pintail/Alberta/121/79).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=11450;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89299445; PubMed=2525836;
RA Treator J.J., Snyder M.H., London W.T., Murphy B.R.;
RA "The B allele of the NS gene of avian influenza viruses, but not the
RT A allele, attenuates a human influenza A virus for squirrel
RT monkeys.";
RT Virology 171:1-9(1989).
RL -1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
ALTERNATIVE SPLICING OF SEGMENT 8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25372; AAA43561.1;
DR PIR: E32663; MNIV17.
DR HSSP: P03495; IAIL.
DR InterPro: IPR000256;
DR Pfam: PF00600; Flu_NS1; 1.
KW Nonstructural protein; Alternative splicing.
SQ SEQUENCE 230 AA; 26120 MW; C938C48FB1ED1BCA CRC64;

Query Match
Best Local Similarity 73.5%; Score 36; DB 1; Length 230;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ISRFAMG 7
1 |||||
198 IORFAMG 204

RESULT 7
VNS1_IAP13 STANDARD; PRT; 230 AA.
VNS1_IAP13
01-JAN-1990 (Rel. 13, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
01-MAY-1992 (Rel. 22, Last annotation update)
DE NONSTRUCTURAL PROTEIN NS1.
OS Influenza A virus (strain A/Pintail/Alberta/358/79).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OX NCBI_TaxID=11452;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89299445; PubMed=2525836;
RA Treator J.J., Snyder M.H., London W.T., Murphy B.R.;
RA "The B allele of the NS gene of avian influenza viruses, but not the
RT A allele, attenuates a human influenza A virus for squirrel
RT monkeys.";
RT Virology 171:1-9(1989).
RL -1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
ALTERNATIVE SPLICING OF SEGMENT 8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE POLY(A) POLYMERASE TYPE 3 (EC 2.7.7.19) (PAP) (POLYNUCLEOTIDE
 DE ADENYLTRANSFERASE) (FRAGMENT)
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 CC Xenopodinae; Xenopus.
 CC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ovary;
 RC MEDLINE=96079940; PubMed=7489490;
 RA Ballantyne S., Bilger A., Astrom J., Virtanen A., Wickens M.;
 RT "Poly (A) polymerases in the nucleus and cytoplasm of frog oocytes:
 RT dynamic changes during oocyte maturation and early development."
 RT RNA 1:64-78(1995).
 CC -1- FUNCTION: POLYMERASE THAT CREATES THE 3' POLY(A) TAIL OF MRNA'S.
 CC MAY ACQUIRE SPECIFICITY THROUGH INTERACTION WITH A CLEAVAGE AND
 CC POLYADENYLATION FACTOR (CPSF)
 CC -1- CATALYTIC ACTIVITY: N ATP + (NUCLEOTIDE)(M) = N PYROPHOSPHATE +
 CC (NUCLEOTIDE)(M+N).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: 019975; AAC59747.1; -
 CC InterPro: IPR002934; -
 CC Pfam: PF01909; NTP_transf_2; 1.
 CC Rfam: Rfam processing; Transferase; Transcription; RNA-binding;
 CC RV Nuclear protein.
 CC FT ACT_SITE 110 110 BY SIMILARITY.
 CC FT ACT_SITE 112 112 BY SIMILARITY.
 CC FT ACT_SITE 164 164 BY SIMILARITY.
 CC FT DOMAIN 382 390 NUCLEAR LOCALIZATION SIGNAL 1 (BY
 CC SIMILARITY).
 CC FT NON_TER 400 400
 CC FT SEQUENCE 400 AA; 46020 MW; AB4347C074E60CA3 CRC64;
 CC -----
 CC Query Match 75.5%; Score 37; DB 1; Length 400;
 CC -st Local Similarity 55.6%; Pred. No. 8.4;
 CC Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC 1 ISRFAMG 9
 CC 376 LAMFAMG 384
 CC -----
 CC RESULT 3
 CC NSL_IACKG
 CC ID NSL_IACKG STANDARD; PRT; 230 AA.
 CC P30909;
 CC 01-JUL-1993 (Rel. 26, Created)
 CC 01-JUL-1993 (Rel. 26, Last sequence update)
 CC 01-JUL-1993 (Rel. 26, Last annotation update)
 CC NONSTRUCTURAL PROTEIN NS1.
 CC S Influenza A virus (strain A/Chicken/Germany/7/49).
 CC C Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 CC C Influenza virus A and B group; Influenza A viruses;
 CC C Influenza A virus.
 CC X NCBI_TaxID=11339;
 CC N [1]
 CC P SEQUENCE FROM N.A.
 CC MEDLINE=91306439; PubMed=1830182;
 CC Ludvig S., Schultz U., Mandler J., Fitch W.M., Scholtissek C.;

RT "Phylogenetic relationship of the nonstructural (NS) genes of
 RT Influenza A viruses".
 RL Virology 183:566-577(1991).
 CC -1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF SEGMENT 8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M55464; AAA3121.1; -
 CC DR HSP; P03495; IAIL.
 CC DR InterPro: IPR000256; -
 CC DR Pfam: PF00600; Flu_NS1; 1.
 CC KW Nonstructural protein; Alternative splicing.
 CC SEQUENCE 230 AA; 26019 MW; 60C643247F62143F CRC64;
 CC -----
 CC Query Match 73.5%; Score 36; DB 1; Length 230;
 CC Best Local Similarity 85.7%; Pred. No. 7.5;
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC 1 ISRFAMG 7
 CC 198 IORFAMG 204
 CC -----
 CC RESULT 4
 CC VNSL_IADA2
 CC ID VNSL_IADA2 STANDARD; PRT; 230 AA.
 CC P03501; P13136;
 CC DT 21-JUL-1986 (Rel. 01, Created)
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
 CC 01-MAY-1992 (Rel. 22, Last annotation update)
 CC DE NONSTRUCTURAL PROTEIN NS1.
 CC OS Influenza A virus (strain A/Duck/Alberta/60/76), and
 CC Influenza A virus (strain A/Mallard/Alberta/82/78).
 CC C Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 CC C Influenza virus A and B group; Influenza A viruses;
 CC C Influenza A virus.
 CC X NCBI_TaxID=11347; 11432;
 CC N [1]
 CC P SEQUENCE FROM N.A.
 CC STRAIN=A/DUCK/ALBERTA/60/76;
 CC MEDLINE=81276929; PubMed=6927848;
 CC Baez M., Zazra J.J., Elliott R.M., Young J.F., Palese P.,
 CC "Nucleotide sequence of the NS1/NS2 overlapping gene structure in a
 CC RNA: conservation of the NS1/NS2 overlapping gene structure in a
 CC divergent influenza virus RNA segment".
 CC Virology 113:397-402(1981).
 CC [2]
 CC RP SEQUENCE FROM N.A.
 CC STRAIN=A/MALLARD/ALBERTA/82/78;
 CC MEDLINE=89299445; PubMed=2525836;
 CC Treanor J.J., Snyder M.H., London W.T., Murphy B.R.;
 CC "The B allele of the NS gene of avian influenza viruses, but not the
 CC A allele, attenuates a human influenza A virus for squirrel
 CC monkeys".
 CC Virology 171:1-9(1989).
 CC -1- ALTERNATIVE PRODUCTS: NS1 AND NS2 PROTEINS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF SEGMENT 8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2001, 16:52:30 ; Search time 12.81 Seconds

(without alignments)
24.067 Million cell updates/sec

Title: US-09-646-579-2

Perfect score: 49

Sequence: 1 ISRFAMGEV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	134	1	DCAC_ACICA
2	37	75.5	400	1	PAP3_XENIA
3	36	73.5	230	1	VNS1_IADK
4	36	73.5	230	1	VNS1_IADK2
5	36	73.5	230	1	VNS1_IADK6
6	36	73.5	230	1	VNS1_IADK1
7	36	73.5	230	1	VNS1_IADK3
8	36	73.5	230	1	VNS1_IADK5
9	36	73.5	230	1	VNS1_IADK7
10	36	73.5	230	1	VNS1_IADK9
11	36	73.5	230	1	VNS1_IADK11
12	36	73.5	230	1	VNS1_IADK13
13	36	73.5	230	1	VNS1_IADK15
14	36	73.5	230	1	VNS1_IADK17
15	36	73.5	230	1	VNS1_IADK19
16	36	73.5	230	1	VNS1_IADK21
17	36	73.5	230	1	VNS1_IADK23
18	36	73.5	230	1	VNS1_IADK25
19	36	73.5	230	1	VNS1_IADK27
20	36	73.5	230	1	VNS1_IADK29
21	36	73.5	230	1	VNS1_IADK31
22	36	73.5	230	1	VNS1_IADK33
23	36	73.5	230	1	VNS1_IADK35
24	36	73.5	230	1	VNS1_IADK37
25	36	73.5	230	1	VNS1_IADK39
26	36	73.5	230	1	VNS1_IADK41
27	36	73.5	230	1	VNS1_IADK43
28	36	73.5	230	1	VNS1_IADK45
29	36	73.5	230	1	VNS1_IADK47
30	36	73.5	230	1	VNS1_IADK49
31	36	73.5	230	1	VNS1_IADK51
32	36	73.5	230	1	VNS1_IADK53
33	36	73.5	230	1	VNS1_IADK55

34	33	67.3	612	1	NUSM_ONCMY	P48176 oncorhynchu
35	33	67.3	612	1	NUSM_SALSA	O952m3 salmo salar
36	33	67.3	613	1	NUSM_POIOR	O95918 polyperus
37	33	67.3	666	1	MGGB_BACCL	P30537 bacillus ca
38	33	65.3	74	1	MGGB_METRA	O32868 methanopyru
39	32	65.3	90	1	MGGB_HAETN	P44048 haemophilus
40	32	65.3	173	1	MGGB_CHICK	P15720 gallus galli
41	32	65.3	257	1	MGGB_ORYSA	P93411 oryza sativ
42	32	65.3	293	1	MGGB_YEAST	P40558 saccharomyc
43	32	65.3	425	1	MGGB_YEAST	P55639 rhizobium s
44	32	65.3	427	1	MGGB_YEAST	P76271 escherichia
45	32	65.3	454	1	MGGB_YEAST	P29434 buchnera ap

ALIGNMENTS

RESULT	ID	DCAC_ACICA	STANDARD	PRT	134 AA
1	AC	P20370:			
DT	01-FEB-1991	(Rel. 17, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	15-JUL-1999	(Rel. 38, Last annotation update)			
DE	4-CARBOXYMUCONOLACTONE DECARBOXYLASE (EC 4.1.1.44) (CMD)				
GN	PCAC				
OS	Acinetobacter calcoaceticus.				
OC	Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;				
OC	Acinetobacter.				
OX	NCBI_TaxID=471;				
FN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BD143 / ADP1;				
RA	MEDLINE=90130333; PubMed=2298704;				
RA	Hartnett C., Neidle E.L., Ngai K.-L., Ornstom L.N.;				
RT	"DNA sequences of genes encoding Acinetobacter calcoaceticus				
RT	protocatalase 3,4-dioxygenase: evidence indicating shuffling of				
RT	genes and of DNA sequences within genes during their evolutionary				
RT	divergence.;"				
RL	J. Bacteriol. 172:956-966(1990).				
CC	-1- CATALYTIC ACTIVITY: 2-CARBOXY-5-OXO-2,5-DIHYDROFURAN-2-ACETATE -				
CC	5-OXO-4,5-DIHYDROFURAN-2-ACETATE + CO(2).				
CC	-1- PATHWAY: THIRD STEP IN THE CATABOLISM OF PROTOCATALASE TO				
CC	SUCCINATE-AND ACETYL-COA IN THE BETA-KETODIAPYRATE PATHWAY.				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC					
DR	EMBL: L05770; AAC37152.1; -				
DR	PIR: B35119; B35119.				
KW	Aromatic hydrocarbons catabolism; Lyase; Decarboxylase.				
SQ	SEQUENCE 134 AA; 15324 MW; 3AEDDELEIC2834A CRC64;				

Query Match 100.0%; Score 49; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMGEV 9
DB 41 ISRFAMGEV 49

RESULT 2
ID PAP3_XENIA STANDARD; PRT; 400 AA.
AC P51006;
DT 01-OCT-1996 (Rel. 34, Created)

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISRFAMG 7

DB 196 IORFAMG 202

RESULT 12

Q900E3 PRELIMINARY; PRT; 229 AA.

AC Q900E3; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE NONSTRUCTURAL PROTEIN 1 (FRAGMENT).

GN NSI.

OS Influenza A virus (A/Turkey/California/189/66(H9N2)).

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;

OC Influenza virus A and B group; Influenza A viruses; Influenza A virus.

OC NCB1_TaxID=97390;

RN NCB1_TaxID=97390;

SEQUENCE FROM N.A.

STRAIN-A/TURKEY/CALIFORNIA/189/66;

RA MEDLINE-99362763; PubMed-10430948;

RA Guan Y., Shortridge K.F., Krauss S., Webster R.G.,

RT Molecular characterization of H9N2 influenza viruses: were they the

RT donors of the 'internal' genes of H5N1 viruses in hong kong?;

RT Proc. Natl. Acad. Sci. U.S.A. 96:9363-9367(1999).

DR EMBL; AF156485; AAC52961.1;

DR HSSP; P03495; IATL.

DR Interpro; IPR000256;

DR Pfam; PF00600; Flu_NSI; 1.

DR Nonstructural protein.

FT NON_TER

SEQUENCE 229 AA; 25964 MW; CAB9B7B86FBC3BD CRC64;

SO

Query Match

Best Local Similarity 73.5%; Score 36; DB 14; Length 229;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISRFAMG 7

DB 197 IORFAMG 203

RESULT 13

Q41651 PRELIMINARY; PRT; 230 AA.

AC Q41651; 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE NONSTRUCTURAL PROTEIN 1.

GN NSI.

OS Influenza A virus.

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;

OC Influenza virus A and B group; Influenza A viruses.

OC NCB1_TaxID=11320;

RN NCB1_TaxID=11320;

SEQUENCE FROM N.A.

STRAIN-A/TURKEY/MINNESOTA/3689-1551/91;

RA Garcia M., Suarez D.L., Crawford J.M., Latimer J.W., Slemmons R.D.,

RA Swayne D.E., Purdie M.L.,

RL Virus Res. 0:0-0(1997).

EMBL; U85378; AAC40657.1;

HSSP; P03495; IATL.

Interpro; IPR000256;

Pfam; PF00600; Flu_NSI; 1.

Prodrom; PD000613;

Nonstructural protein.

SEQUENCE 230 AA; 26199 MW; 38071EC8560ACEF3 CRC64;

Query Match

Best Local Similarity 73.5%; Score 36; DB 14; Length 230;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISRFAMG 7

DB 198 IORFAMG 204

RESULT 14

Q41653 PRELIMINARY; PRT; 230 AA.

AC Q41653; 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE NONSTRUCTURAL PROTEIN 1.

GN NSI.

OS Influenza A virus.

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;

OC Influenza virus A and B group; Influenza A viruses.

OC NCB1_TaxID=11320;

RN NCB1_TaxID=11320;

SEQUENCE FROM N.A.

STRAIN-A/DUCK/MICHIGAN/80;

RA Garcia M., Suarez D.L., Crawford J.M., Latimer J.W., Slemmons R.D.,

RA Swayne D.E., Purdie M.L.,

RL Virus Res. 0:0-0(1997).

EMBL; U85381; AAC40663.1;

HSSP; P03495; IATL.

Interpro; IPR000256;

Pfam; PF00600; Flu_NSI; 1.

Prodrom; PD000613;

Nonstructural protein.

SEQUENCE 230 AA; 26067 MW; 38A85BD8E1C5CF96 CRC64;

SO

Query Match

Best Local Similarity 73.5%; Score 36; DB 14; Length 230;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISRFAMG 7

DB 198 IORFAMG 204

RESULT 15

Q41654 PRELIMINARY; PRT; 230 AA.

AC Q41654; 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE NONSTRUCTURAL PROTEIN 1.

GN NSI.

OS Influenza A virus.

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;

OC Influenza virus A and B group; Influenza A viruses.

OC NCB1_TaxID=11320;

RN NCB1_TaxID=11320;

SEQUENCE FROM N.A.

STRAIN-A/TURKEY/MINNESOTA/3689-1551/91;

RA Garcia M., Suarez D.L., Crawford J.M., Latimer J.W., Slemmons R.D.,

RA Swayne D.E., Purdie M.L.,

RL Virus Res. 0:0-0(1997).

EMBL; U85382; AAC40665.1;

HSSP; P03495; IATL.

Interpro; IPR000256;

Pfam; PF00600; Flu_NSI; 1.

Prodrom; PD000613;

Nonstructural protein.

SEQUENCE 230 AA; 25920 MW; 4381024D68E06D28 CRC64;

Query Match 73.5%; Score 36; DB 14; Length 230;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ISRFAMG 7
1 11111
b 198 IORFAMG 204

Search completed: October 18, 2001, 16:53:11
Job time: 156 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2001, 16:50:35 ; Search time 32.41 Seconds
(without alignments)
14.964 Million cell updates/sec

Title: US-09-646-579-1

Perfect score: 45
Sequence: 1 RFSAMGAE 8

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

num DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	8	20	AAV42321
2	36	80.0	500	20	AAV04867
3	36	80.0	527	20	AAV04868
4	34	75.6	25	21	AAV90994
5	34	75.6	343	21	AAV73418
6	34	75.6	522	21	AAV58875
7	34	75.6	544	21	AAV58874
8	34	75.6	578	21	AAV58876
9	34	75.6	635	21	AAV58877
10	34	75.6	659	21	AAV58871
11	34	75.6	705	21	AAV58872

12	34	75.6	705	21	AAV58873	Human peptidase NA
13	34	75.6	740	21	AAV58870	Human peptidase NA
14	33	73.3	8	21	AAV12947	Helicobacterium NA
15	33	73.3	16	16	AAV71029	HSV-2 protease sub
16	33	73.3	19	18	AAV37707	C-terminal additio
17	33	73.3	20	14	AAV32141	HSV-1 enzyme relea
18	33	73.3	55	20	AAV28579	Secreted peptide c
19	33	73.3	106	21	AAV54338	Human pancreatic c
20	33	73.3	201	22	AAV79733	Corynebacterium q1
21	33	73.3	246	22	AAV79732	Corynebacterium q1
22	33	73.3	287	13	AAV28644	U26 protease dele
23	33	73.3	306	17	AAV88559	Herpes simplex vir
24	33	73.3	314	13	AAV29623	Tet e gene product
25	33	73.3	314	13	AAV37873	Tet e gene product
26	33	73.3	314	13	AAV34544	Tet e gene product
27	33	73.3	350	18	AAV09968	HSV-1 (F) protease
28	33	73.3	350	18	AAV09978	HSV-1 (F) protease
29	33	73.3	350	18	AAV09977	HSV-1 (F) protease
30	33	73.3	350	18	AAV09976	HSV-1 (F) protease
31	33	73.3	350	18	AAV09975	HSV-1 (F) protease
32	33	73.3	350	18	AAV09974	HSV-1 (F) protease
33	33	73.3	350	18	AAV09973	HSV-1 (F) protease
34	33	73.3	350	18	AAV09972	HSV-1 (F) protease
35	33	73.3	350	18	AAV09971	HSV-1 (F) protease
36	33	73.3	350	18	AAV09970	HSV-1 (F) protease
37	33	73.3	350	18	AAV09969	HSV-1 (F) protease
38	33	73.3	350	18	AAV09967	HSV-1 (F) protease
39	33	73.3	350	18	AAV09966	HSV-1 (F) protease
40	33	73.3	350	18	AAV09965	HSV-1 (F) protease
41	33	73.3	350	18	AAV09964	HSV-1 (F) protease
42	33	73.3	350	18	AAV09963	HSV-1 (F) protease
43	33	73.3	350	18	AAV09962	HSV-1 (F) protease
44	33	73.3	350	18	AAV09961	HSV-1 (F) protease
45	33	73.3	350	18	AAV09960	HSV-1 (F) protease

ALIGNMENTS

RESULT 1
AAV42321
ID AAV42321 standard; peptide; 8 AA.
XX
AC AAV42321;
XX
DT 06-DEC-1999 (first entry)
XX
DE Acinetobacter-derived peptide epitope #1.
XX
KW Acinetobacter; myelin; spongiform encephalopathy; CJD; Immune response;
KW Creutzfeldt-Jakob disease; multiple sclerosis; antibody; epitope;
KW detection.
XX
OS Synthetic.
OS Acinetobacter sp.
XX
PN MO9947932-A2.
XX
PD 23-SEP-1999.
XX
PF 19-MAR-1999; 99WO-GB00876.
XX
PR 19-MAR-1998; 98GB-0005913.
XX
PA (UNIV) UNIV LONDON KING'S COLLEGE.
XX
PI Edinger A;
XX
XX WPI; 1999-571874/48.
XX
XX Detecting de-myelinating disease or spongiform encephalopathy
XX
PS Claim 12; Page 7; 11pp; English.

This sequence represents a synthetic peptide epitope (#1) derived from species of *Acinetobacter*, which is sufficiently conformationally similar to natural *Acinetobacter* peptide epitopes to bind the corresponding antibodies. *Acinetobacter* species may be implicated in spongiform encephalopathies such as Creutzfeldt-Jakob disease or in demyelinating diseases such as multiple sclerosis. Involvement of *Acinetobacter* forms the basis for a model for the development of de-myelinating diseases (and an alternative model for the development of spongiform encephalopathies). This model involves the phenomenon of molecular mimicry in which mammals exposed to certain bacteria containing peptide sequences which mimic myelin peptides, experience an autoimmune response. A novel method for detecting a de-myelinating disease or spongiform encephalopathy in mammals has been developed which comprises testing a biological sample obtained from the mammal for immunoglobulin A (IgA) antibodies indicative of infection by *Acinetobacter* species. This sequence can be used as a test antigen in a kit to detect such antibodies. The method is useful for diagnosing de-myelinating diseases and spongiform encephalopathies in animals and humans, especially bovine spongiform encephalopathy, multiple sclerosis or Creutzfeldt-Jacob disease in humans. The method is useful for the early detection of these infections, resulting in early treatments.

Sequence 8 AA;

Query Match 100.0%; Score 45; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 RFSAMGAE 8
|||||||
1 rfawgae 8

AY04867 standard; Protein: 500 AA.

AY04867;

06-JUL-1999 (first entry)

Mycobacterium species protein sequence 19D.

Secreted protein; Mycobacterium; primer; PCR; amplification; probe; hybridisation; detection; vaccine; immunisation; infection.

Mycobacterium sp.

W09909186-A2.

25-FEB-1999.

14-AUG-1998; 98MO-FR01813.

11-SEP-1997; 97FR-0011325.

14-AUG-1997; 97FR-0010404.

(INSP) INST PASTEUR.

Gicquel B, Lim EM, Pelletier V, Portnoi D, Goguet de la Salmoniere Y; Guigneno A;

WPI; 1999-181045/15.

N-PSDB; AAX34119.

Mycobacterial DNA vectors containing reporter constructs - for identifying coding or promoter sequences involved in infection-associated protein expression

Claim 32; Fig 19D; 309pp; French.

Sequences AY04742-Y05000 and AY07201-Y07204 represent secreted proteins from various *Mycobacterium* species microorganisms. The encoding nucleotide sequences can be used as primers and probes for methods for detecting and identifying mycobacteria, especially belonging to the *M. tuberculosis* complex. The encoded proteins can be used in vaccines for immunisation against a bacterial or viral infection.

Sequence 500 AA;

Query Match 80.0%; Score 36; DB 20; Length 500;
Best Local Similarity 75.0%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RFSAMGAE 8
|||||
Db 309 rfawgae 316

AY04868 standard; Protein: 527 AA.

AY04868;

06-JUL-1999 (first entry)

Mycobacterium species protein sequence 19F.

Secreted protein; Mycobacterium; primer; PCR; amplification; probe; hybridisation; detection; vaccine; immunisation; infection.

Mycobacterium sp.

W09909186-A2.

25-FEB-1999.

14-AUG-1998; 98MO-FR01813.

11-SEP-1997; 97FR-0011325.

14-AUG-1997; 97FR-0010404.

(INSP) INST PASTEUR.

Gicquel B, Lim EM, Pelletier V, Portnoi D, Goguet de la Salmoniere Y; Guigneno A;

WPI; 1999-181045/15.

N-PSDB; AAX34120.

Mycobacterial DNA vectors containing reporter constructs - for identifying coding or promoter sequences involved in infection-associated protein expression

Claim 32; Fig 19F; 309pp; French.

Sequences AY04742-Y05000 and AY07201-Y07204 represent secreted proteins from various *Mycobacterium* species microorganisms. The encoding nucleotide sequences can be used as primers and probes for methods for detecting and identifying mycobacteria, especially belonging to the *M. tuberculosis* complex. The encoded proteins can be used in vaccines for immunisation against a bacterial or viral infection.

Sequence 527 AA;

Query Match 80.0%; Score 36; DB 20; Length 527;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RFSAMGAE 8
|||||

Db 336 rfafwgae 343

RESULT 4

ID AAY90994 standard; Peptide; 25 AA.

AC AAY90994;

DT 05-SEP-2000 (first entry)

DE Inhibitory clone p+/1 peptide sequence SEQ ID NO:100.

XX Escherichia coli; E. coli; randomised peptide library; identification;
 XX stabilised bioactive peptide; synthesis; intracellular selection;
 KW screening; lac operon; protease resistant; peptidase resistant;
 KW ROP protein; glutathione sulphotransferase; thiorodoxin; infection;
 KW maltose binding protein; glutathione reductase; antimicrobial;
 KW antibacterial.

OS Escherichia coli.

XX Synthetic.

PN W0200022112-A1.

PD 20-APR-2000.

PE 12-OCT-1999; 99WO-US23731.

XX 13-OCT-1998; 98US-0104013.

PR 14-DEC-1998; 98US-0112150.

PA (UTGE-) UNIV GEORGIA RES FOUND INC.

XX (ALTMAN) ALTMAN E.

PI Altman E;

XX WPI; 2000-317972/27.
 XX Identifying recombinantly an antimicrobial bioactive peptide used as a
 XX therapeutic agent involves transforming a host cell with expression
 XX vector with slightly regulable control region and measuring its
 XX inhibition

XX Example 3; Page 77; 135pp; English.

XX The present invention describes a method for identifying a bioactive
 XX peptide (BP) involving transforming a cell with an expression vector
 XX comprising a tightly regulable control region operably linked to a
 XX nucleic acid sequence encoding a peptide (P), growing the transformed
 XX cell under conditions that repress expression of (P) and then inducing
 XX its expression which, if is inhibitory to host cell growth, is
 XX indicative of BP expression. An antimicrobial peptide from the present
 XX invention, which is stabilised, is used for treating a patient having a
 XX condition treatable with a peptide drug. The stabilised peptides are
 XX also used for inhibiting the growth of a microbe. The new antibacterial
 XX peptides are useful to treat various pathogenic bacteria such as
 XX Staphylococci, Streptococci and Enterococci which are the primary causes
 XX of nosocomial infections. Novel inhibitor peptides identified by the
 XX method can be medical treatments and therapies directed against
 XX microbial infection. Also, these novel inhibitor peptides can be used,
 XX in turn, to identify additional novel antibacterial peptides using a
 XX synthetic approach, and can also be used to elucidate potential new drug
 XX targets. The inhibitor peptide target which is inactivated is identified
 XX using reverse genetics by isolating mutants that are no longer inhibited
 XX by the peptide. These mutants are then mapped in order to precisely
 XX determine the protein target that is inhibited. AAY56033 to AAY56106 and
 XX AAY90964 to AAY90999 are sequences used in the exemplification of the
 XX present invention.

XX Sequence 25 AA;

Query Match 75.6%; Score 34; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 9;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SAWGAE 8

IIIIII

DB 9 sawgae 14

RESULT 5

ID AAY73418 standard; Protein; 343 AA.

AC AAY73418;

DT 29-FEB-2000 (first entry)

DE Human secreted protein clone yc48.1 protein sequence SEQ ID NO:58.

XX Human; secreted protein; immunostimulatory; haemostatic; cytokine;

XX Proliferative; differentiative; chemotactic; chemokinetic; vaccine;

XX thrombolytic; antiinflammatory; cytostatic; immunosuppressive;

XX gene therapy.

XX Homo sapiens.

XX W09958642-A2.

PN 18-NOV-1999.

PD 14-MAY-1999; 99WO-US10843.

XX 14-MAY-1998; 98US-0085472.

PR 17-AUG-1998; 98US-0096824.

PR 11-SEP-1998; 98US-0099843.

PR 11-SEP-1998; 98US-0099950.

PR 15-SEP-1998; 98US-0100424.

PR 29-SEP-1998; 98US-0102329.

PR 09-OCT-1998; 98US-0103615.

PR 11-DEC-1998; 98US-0111799.

PR 14-DEC-1998; 98US-0112159.

PR 31-DEC-1998; 98US-0114415.

PR 10-FEB-1999; 98US-0248059.

PR 06-APR-1999; 99US-0287150.

PR 13-MAY-1999; 99US-0311021.

XX (GEMY) GENETICS INST INC.

XX Wong GG, Clark HF, Fecthel K, Agostino MJ;

XX WPI; 2000-053095/04.

XX N-PSDB; AA252503.

XX Novel polynucleotides and proteins having biological activities which

XX make them suitable for treating, preventing or ameliorating medical

XX conditions in humans or animals -

XX Claim 67; Page 621-622; 730pp; English.

XX The present invention describes human secreted proteins encoded by
 XX polynucleotides obtained from adult testes, foetal brain, adult brain,
 XX brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
 XX cDNA libraries. The polynucleotides and proteins are predicted to have
 XX biological activities which would make them suitable for treating,
 XX preventing or ameliorating medical conditions in humans and animals.
 XX Suggested activities include nutritional activity, cytokine and cell
 XX proliferation/differentiation activity, immune stimulating (e.g. as
 XX vaccines) or suppressing activity, haematopoiesis regulating activity,
 XX tissue growth activity, activin/inhibin activity, chemotactic/
 XX chemokinetic activity, haemostatic and thrombolytic activity, receptor/
 XX ligand activity, anti-inflammatory activity, cadherin/tumour invasion
 XX suppressor activity, and tumour inhibition activity. The polynucleotides
 XX are also stated to be useful for gene therapy. Therapeutic compositions

are also presently valuable for veterinary applications. AA252475 to AA252581 encode human secreted proteins, and AA73390 to AA73500 represent human secreted proteins, given in the present invention.

Sequence 343 AA:

Query Match 75.6%; Score 34; DB 21; Length 343;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 FRSAMGA 7
111111
136 fisaaga 142

ESULT 6
AY58875
AA58875 standard; Protein: 522 AA.

AA58875;

08-MAY-2000 (first entry)

Human peptidase NALAD-ase L splice variant (premature termination).

NALAD-ase L; N-acetylated alpha-linked acidic dipeptidase; human;
chromosome 11q12; prostate cancer; neurodegenerative disease;
Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
peripheral neuropathy; Huntington's disease; acute brain injury;
multiple sclerosis; peripheral nerve trauma; ischemia; dementia;
gene therapy; diagnosis; neurotropic; neuroprotective; neuroleptic;
antiparkinsonian; anticonvulsant; vasotropic; splice variant.

Homo sapiens.

Key Location/Qualifiers
6..27

Domain /note= "membrane-spanning domain"

Misc-difference 100 /note= "encoded by TMC"

WO200004157-A2.

27-JAN-2000.

14-JUL-1999; 99WO-GB02241.

14-JUL-1998; 98GB-0015284.

(JANC) JANSSEN PHARM NV.

Pangalos M, Neefs JEFM, Peeters DCG;

WPI: 2000-182424/16.

N-PSDB; AA258309.

New human N-acetylated alpha-linked acidic dipeptidases for treating neural disorders e.g. Alzheimer's disease, schizophrenia and Parkinson's disease

Claim 11; Fig -; 95pp; English.

The present sequence is that of a human N-acetylated alpha-linked acidic dipeptidase L (NALAD-ase L) splice variant, predicted from cDNA (see AA258309) obtained during 3' RACE amplifications of small intestine, colon, brain and foetal brain cDNA. The splice variant is prematurely terminated compared with the active protein (see AA258870) owing to a 173-nucleotide deletion that results in a frame-shift. Several NALAD-ase L splice variants (see AA258871-77) have been identified. Their biological significance has yet to be determined, but inactive variants may be used to regulate the levels of active protein. The invention provides human NALAD-ase

CC L (and splice variant), II and IV polypeptides, cDNAs, antisense
CC nucleic acids, vectors, host cells, transgenic organisms, antagonists
CC and agonists. These are useful for treating neural disorders such as
CC Alzheimer's disease, schizophrenia, ALS, Parkinson's disease,
CC peripheral neuropathy, Huntington's disease, acute brain injury,
CC multiple sclerosis, exposure to neurotoxins, peripheral nerve trauma,
CC ischemia or dementia (claimed).
CC Note: The present sequence is not shown in the specification but is
CC derived from the NALAD-ase L sequence given in figure 1 and
CC figure 3.
XX
SQ Sequence 522 AA;

Query Match 75.6%; Score 34; DB 21; Length 522;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 FRSAMGA 8
111111
DB 410 fawgae 416

RESULT 7
AA58874
ID AA58874 standard; Protein: 544 AA.

AC AA58874;

08-MAY-2000 (first entry)

Human peptidase NALAD-ase L splice variant (premature termination).

NALAD-ase L; N-acetylated alpha-linked acidic dipeptidase; human;
chromosome 11q12; prostate cancer; neurodegenerative disease;
Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
peripheral neuropathy; Huntington's disease; acute brain injury;
multiple sclerosis; peripheral nerve trauma; ischemia; dementia;
gene therapy; diagnosis; neurotropic; neuroprotective; neuroleptic;
antiparkinsonian; anticonvulsant; vasotropic; splice variant.

Homo sapiens.

Key Location/Qualifiers
6..27

Domain /note= "membrane-spanning domain"

Misc-difference 100 /note= "encoded by TMC"

WO200004157-A2.

27-JAN-2000.

14-JUL-1999; 99WO-GB02241.

14-JUL-1998; 98GB-0015284.

(JANC) JANSSEN PHARM NV.

Pangalos M, Neefs JEFM, Peeters DCG;

WPI: 2000-182424/16.

N-PSDB; AA258308.

New human N-acetylated alpha-linked acidic dipeptidases for treating neural disorders e.g. Alzheimer's disease, schizophrenia and Parkinson's disease

Claim 11; Fig -; 95pp; English.

The present sequence is that of a human N-acetylated alpha-linked acidic dipeptidase L (NALAD-ase L) splice variant, predicted from cDNA (see AA258308) obtained during 3' RACE amplifications of small

Intestine, colon, brain and foetal brain cDNA. The splice variant is prematurely terminated compared with the active protein (see AAY58870) owing to a 91 nucleotide deletion that results in a frame-shift. Several NALAD-ase L splice variants (see AAY58871-77) have been identified. Their biological significance has yet to be determined, but inactive variants may be used to regulate the levels of active protein. The invention provides human NALAD-ase L (and splice variant), II and IV polypeptides, cDNAs, antisense nucleic acids, vectors, host cells, transgenic organisms, antagonists and agonists. These are useful for treating neural disorders such as Alzheimer's disease, schizophrenia, ALS, Parkinson's disease, peripheral neuropathy, Huntington's disease, acute brain injury, multiple sclerosis, exposure to neurotoxins, peripheral nerve trauma, ischemia or dementia (claimed).

Note: The present sequence is not shown in the specification but is derived from the NALAD-ase L sequence given in figure 1 and figure 3.

Sequence 544 AA:

Very Match 75.6%; Score 34; DB 21; Length 544;
 at Local Similarity 71.4%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

2 FSAMGAE 8
 1:|||||
 DB 410 faswgae 416

RESULT 8
 AAY58876
 ID AAY58876 standard; Protein: 578 AA.

AA58876;
 08-MAY-2000 (first entry)

Human peptidase NALAD-ase L splice variant (premature termination).
 NALAD-ase L: N-acetylated alpha-linked acidic dipeptidase; human;
 chromosome 11q12; prostate cancer; neurodegenerative disease;
 Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
 peripheral neuropathy; Huntington's disease; acute brain injury;
 multiple sclerosis; peripheral nerve trauma; ischemia; dementia;
 gene therapy; diagnosis; nootropic; neuroprotective; neuroleptic;
 antiparkinsonian; anticonvulsant; vasotropic; splice variant.

Homo sapiens.

Key Location/Qualifiers
 Domain 6..27
 /note= "membrane-spanning domain"
 Misc-difference 100
 /note= "encoded by TNC"

MO200004157-A2.

27-JAN-2000.

14-JUL-1999; 99WO-GB02241.

14-JUL-1998; 98GB-0015284.

(JANC) JANSSEN PHARM NV.

Pangalos M, Neefs JEFM, Peeters DCG;

WPI: 2000-182424/16.

N-PSDB; AA258310.

New human N-acetylated alpha-linked acidic dipeptidases for treating neural disorders e.g. Alzheimer's disease, schizophrenia and

Parkinson's disease
 Claim 11: Fig -; 95pp; English.

The present sequence is that of a human N-acetylated alpha-linked acidic dipeptidase L (NALAD-ase L) splice variant, predicted from cDNA (see AA258310) obtained during 3' RACE amplifications of small intestine, colon, brain and foetal brain cDNA. The splice variant is prematurely terminated compared with the active protein (see AAY58870) owing to an intronic insertion into cDNA that caused a frame-shift. Several NALAD-ase L splice variants (see AAY58871-77) have been identified. Their biological significance has yet to be determined, but inactive variants may be used to regulate the levels of active protein. The invention provides human NALAD-ase L (and splice variant), II and IV polypeptides, cDNAs, antisense nucleic acids, vectors, host cells, transgenic organisms, antagonists and agonists. These are useful for treating neural disorders such as Alzheimer's disease, schizophrenia, ALS, Parkinson's disease, peripheral neuropathy, Huntington's disease, acute brain injury, multiple sclerosis, exposure to neurotoxins, peripheral nerve trauma, ischemia or dementia (claimed).

Note: The present sequence is not shown in the specification but is derived from the NALAD-ase L sequence given in figure 1 and figure 3.

Sequence 578 AA:

Query Match 75.6%; Score 34; DB 21; Length 578;
 Best Local Similarity 71.4%; Pred. No. 2.2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

2 FSAMGAE 8
 1:|||||
 DB 410 faswgae 416

RESULT 9
 AAY58877
 ID AAY58877 standard; Protein: 635 AA.

AA58877;

08-MAY-2000 (first entry)

Human peptidase NALAD-ase L splice variant (premature termination).
 NALAD-ase L: N-acetylated alpha-linked acidic dipeptidase; human;
 chromosome 11q12; prostate cancer; neurodegenerative disease;
 Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
 peripheral neuropathy; Huntington's disease; acute brain injury;
 multiple sclerosis; peripheral nerve trauma; ischemia; dementia;
 gene therapy; diagnosis; nootropic; neuroprotective; neuroleptic;
 antiparkinsonian; anticonvulsant; vasotropic; splice variant.

Homo sapiens.

Key Location/Qualifiers
 Domain 6..27
 /note= "membrane-spanning domain"
 Misc-difference 100
 /note= "encoded by TNC"

MO200004157-A2.

27-JAN-2000.

14-JUL-1999; 99WO-GB02241.

14-JUL-1998; 98GB-0015284.

(JANC) JANSSEN PHARM NV.

Pangalos M, Neefs JEFM, Peeters DCG;
WPI: 2000-182424/16.
N-PSDB: AA258311.

New human N-acetylated alpha-linked acidic dipeptidases for treating
neural disorders e.g. Alzheimer's disease, schizophrenia and
Parkinson's disease

Claim 11; Fig -; 95pp; English.

The present sequence is that of a human N-acetylated alpha-linked
acidic dipeptidase L (NALAD-ase L) splice variant, predicted from
cDNA (see AA258311) obtained during 3' RACE amplifications of small
intestine, colon, brain and foetal brain cDNA. The splice variant
is prematurely terminated compared with the active protein (see
AA258870) owing to an intronic insertion into cDNA that caused a
frame-shift. Several NALAD-ase L splice variants (see AA258871-77)
have been identified. Their biological significance has yet to be
determined, but inactive variants may be used to regulate the
levels of active protein. The invention provides human NALAD-ase
L (and splice variant), II and IV polypeptides, cDNAs, antisense
nucleic acids, vectors, host cells, transgenic organisms, antagonists
and agonists. These are useful for treating neural disorders such as
Alzheimer's disease, schizophrenia, ALS, Parkinson's disease,
peripheral neuropathy, Huntington's disease, acute brain injury,
multiple sclerosis, exposure to neurotoxins, peripheral nerve trauma,
ischaemia or dementia (claimed).
Note: The present sequence is not shown in the specification but is
derived from the NALAD-ase L sequence given in figure 1 and
figure 3.

Sequence 635 AA;

Query Match 75.6%; Score 34; DB 21; Length 635;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

2 FSAMGAE 8
1:|||||
410 faswgae 416

RESULT 10
AA258871
AA258871 standard; Protein: 699 AA.

AA258871;

08-MAY-2000 (first entry)

Human peptidase NALAD-ase L splice variant (deletion).

NALAD-ase L, N-acetylated alpha-linked acidic dipeptidase; human;
Chromosome 11q12; prostate cancer; neurodegenerative disease;
Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
peripheral neuropathy; Huntington's disease; acute brain injury;
multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;
gene therapy; diagnosis; nootropic; neuroprotective; neuroleptic;
antiparkinsonian; anticonvulsant; vasotropic; splice variant.

Homo sapiens.

Key Location/Qualifiers

Domain

Misc-difference 100 /note="membrane-spanning domain"

/note="encoded by TNC"

MO200004157-A2.

27-JAN-2000.

XX 14-JUL-1999; 99WO-GB02241.
XX 14-JUL-1998; 98GB-0015284.

XX (JANC) JANSSEN PHARM NV.

XX Pangalos M, Neefs JEFM, Peeters DCG;

XX WPI: 2000-182424/16.
XX N-PSDB: AA258305.

XX New human N-acetylated alpha-linked acidic dipeptidases for treating
XX neural disorders e.g. Alzheimer's disease, schizophrenia and
XX Parkinson's disease

XX Claim 11; Fig -; 95pp; English.

XX The present sequence is that of a human N-acetylated alpha-linked
XX acidic dipeptidase L (NALAD-ase L) splice variant, predicted from
XX cDNA (see AA258305) obtained during 5' RACE amplifications of small
XX intestine and colon cDNA. The splice variant carries a deletion of
XX amino acids 161-201 of NALAD-ase L (see AA258870). Several
XX NALAD-ase L splice variants (see AA258871-77) have been identified.
XX Their biological significance has yet to be determined, but spliced
XX out residues may affect levels of glycosylation and the conformation
XX and activity of the protein. The invention provides human NALAD-ase
XX L (and splice variant), II and IV polypeptides, cDNAs, antisense
XX nucleic acids, vectors, host cells, transgenic organisms, antagonists
XX and agonists. These are useful for treating neural disorders such as
XX Alzheimer's disease, schizophrenia, ALS, Parkinson's disease,
XX peripheral neuropathy, Huntington's disease, acute brain injury,
XX multiple sclerosis, exposure to neurotoxins, peripheral nerve trauma,
XX ischaemia or dementia (claimed).
XX Note: The present sequence is not shown in the specification but is
XX derived from the NALAD-ase L sequence given in figure 1 and
XX figure 3.

XX Sequence 699 AA;

Query Match 75.6%; Score 34; DB 21; Length 699;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 FSAMGAE 8
1:|||||
Db 369 faswgae 375

RESULT 11
AA258872

ID AA258872 standard; Protein: 705 AA.

AA258872;

08-MAY-2000 (first entry)

Human peptidase NALAD-ase L splice variant (deletion).

NALAD-ase L, N-acetylated alpha-linked acidic dipeptidase; human;
Chromosome 11q12; prostate cancer; neurodegenerative disease;
Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
peripheral neuropathy; Huntington's disease; acute brain injury;
multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;
gene therapy; diagnosis; nootropic; neuroprotective; neuroleptic;
antiparkinsonian; anticonvulsant; vasotropic; splice variant.

Homo sapiens.

Key Location/Qualifiers

Domain

/note="membrane-spanning domain"

FT Misc-difference 100 /note= "encoded by TMC"
 FT
 FT xx
 PN WO200004157-A2.
 PD 27-JAN-2000.
 PD xx
 PP 14-JUL-1999; 99WO-CB02241.
 PP xx
 PR 14-JUL-1998; 98GB-0015284.
 PR xx
 PA (JANC) JANSSEN PHARM NV.
 PA xx
 P1 Pangalos M, Neefs JEFM, Peeters DCG;
 P1 xx
 DR WPI: 2000-182424/16.
 DR N-PSDB; AAZ58306.
 DR xx
 PT New human N-acetylated alpha-linked acidic dipeptidases for treating
 PT neural disorders e.g. Alzheimer's disease, schizophrenia and
 PT Parkinson's disease -

Claim 11; Fig -; 95pp; English.

The present sequence is that of a human N-acetylated alpha-linked acidic dipeptidase L (MALAD-ase L) splice variant, predicted from cDNA (see AY258307) obtained during 5' RACE amplifications of small intestine and colon cDNA. The splice variant carries a deletion of amino acids 296-301 of MALAD-ase L (see AY258870). Several MALAD-ase L splice variants (see AY258871-77) have been identified. Their biological significance has yet to be determined, but spliced out residues may affect levels of glycosylation and the conformation and activity of the protein. The invention provides human MALAD-ase L (and splice variant), II and IV polypeptides, cDNAs, antisense nucleic acids, vectors, host cells, transgenic organisms, antagonists and agonists. These are useful for treating neural disorders such as Alzheimer's disease, schizophrenia, ALS, Parkinson's disease, peripheral neuropathy, Huntington's disease, acute brain injury, multiple sclerosis, exposure to neurotoxins, peripheral nerve trauma, ischaemia or dementia (claimed).

Note: The present sequence is not shown in the specification but is derived from the MALAD-ase L sequence given in figure 1 and figure 3.

Note: The present sequence is not shown in the specification but is derived from the NALAD-ase L sequence given in figure 1 and figure 3.

SQ Sequence 705 AA;

Query Match	75.68;	Score 34;	DB 21;	length 705;
- Best Local Similarity	71.48;	Pred. No. 2.7e+02;		
itches	5;	Conservative	2;	Mismatches 0; Indels 0; Caps 0

```

OY      2 FSAWGAE 8
         1::1111
db      375 faswgae 381

```

RESULT	12
AAV58873	
ID	AAV58873 standard; Protein; 705 AA

AY58873 ;

08-MAY-2000 (first entry)

Human peptidase NAALAD-ase L splice variant (deletion).

CXX NAAALDase L₂; N-acetylated alpha-linked acidic dipeptidase; human
 CXX chromosome 11q12; prostate cancer; neurodegenerative disease;
 CXX Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
 CXX peripheral neuropathy; Huntington's disease; acute brain injury;
 CXX multiple sclerosis; peripheral nerve trauma; ischemia; dementia;
 CXX gene therapy; diagnosis; neurotropic; neuroprotective; neuroleptic;
 CXX antiparkinsonian; anticonvulsant; vasotropic; splice variant

XX	Homo sapiens.
OS	
XX	
EH	Key
FT	Domain
FT	Location/Qualifiers
FT	6..27
FT	/note- "membrane-spanning domain
FT	Misc-difference 100
XX	/note- "encoded by TNC"

/note- "encoded by TNC"

W0200004157-A2.

27-JAN-2000.

14-JUL-1999; 99WO-GB02241

14-JUL-1998; 98GB-0015284

(JANC) JANSSEN PHARM NV

Pangalos M, Neefs JEFM, Peeters DCG;

WPL; 2000-182424/16
N-PSDB: AA758307

N-PSDB; AA258307

New human N-acetylated alpha-linked acidic dipeptidases for treating neural disorders e.g. Alzheimer's disease, schizophrenia and Parkinson's disease -

Claim 11; Fig -; 95pp; English.

The present sequence is that of a human N-acetylated alpha-linked acidic dipeptidase L (NALAD-ase L) splice variant, predicted from cDNA (see A4258307) obtained during 5' RACE amplifications of small intestine and colon cDNA. The splice variant carries a deletion of amino acids 296-301 of NALAD-ase L (see A4538870). Several NALAD-ase L splice variants (see A4538871-77) have been identified. Their biological significance has yet to be determined, but spliced out residues may affect levels of glycosylation and the conformation and activity of the protein. The invention provides human NALAD-ase L and splice variant), II and IV polypeptides, cDNAs, antisense nucleic acids, vectors, host cells, transgenic organisms, antagonists and agonists. These are useful for treating neural disorders such as Alzheimer's disease, schizophrenia, ALS, Parkinson's disease, peripheral neuropathy, Huntington's disease, acute brain injury, multiple sclerosis, exposure to neurotoxins, peripheral nerve trauma, ischemia or dementia (claimed).

Note: The present sequence is not shown in the specification but is derived from the NALAD-ase L sequence given in Figure 1 and Figure 3.

CC note: the present sequence is not shown in the specification but is
CC derived from the NMAAD-ase L sequence given in figure 1 and
CC figure 3.

Sequence 705 AA;

Query Match	75.68;	Score 34;	DB 21;	Length 705;
Best Local Similarity	71.48;	Pred. No. 2.7e+02;		
Matches	5;	Conservative	0;	Indels 0;
		Mismatches	0;	Gaps 0

QY	2	FSAWGAE	8
		1::1111	
Db	375	faswgae	381

RESULT	13
AAV58870	
ID	AAV58870 standard; Protein; 740 AA
rv	

AC AAY58870;
XX

DT 08-MAY-2000 (first entry)
XX

Human peptidase NAALAD-ase L

KW NAALAD-ase L; N-acetylated alpha-linked dipeptidase; human;

chromosome 11q12; prostate cancer; neurodegenerative disease;
 Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
 peripheral neuropathy; Huntington's disease; acute brain injury;
 multiple sclerosis; peripheral nerve trauma; ischemia; dementia;
 therapy; diagnosis; neurotropic; neuroprotective; neuroleptic;
 antiparkinsonian; anticonvulsant; vasotropic.

Homo sapiens.

Key Location/Qualifiers

Domain 6..27

Misc-difference 100 /note= "membrane-spanning domain"

Modified-site 136 /note= "encoded by TNC"

Modified-site 274 /note= "N-glycosylated"

Modified-site 298 /note= "N-glycosylated"

Modified-site 334 /note= "N-glycosylated"

Modified-site 345 /note= "N-glycosylated"

Modified-site 451 /note= "N-glycosylated"

Modified-site 492 /note= "N-glycosylated"

Modified-site /note= "N-glycosylated"

WO200004157-A2.

27-JAN-2000.

14-JUL-1999; 99WO-GB02241.

14-JUL-1998; 98GB-0015284.

(JANC) JANSSEN PHARM NV.

Pangalos M, Neefs JEFM, Peeters DCG;

WPI: 2000-182424/16.

N-PEDB; AA258304.

New human N-acetylated alpha-linked acidic dipeptidases for treating
 neural disorders e.g. Alzheimer's disease, schizophrenia and
 Parkinson's disease

Claim 10; Fig 1; 95pp; English.

The present sequence is that predicted for human N-acetylated
 alpha-linked acidic dipeptidase L (NALAD-ase L) on the basis of
 isolated cDNA (see AA258304). NALAD-ase L is predicted to be a
 type II integral membrane protein with 78% identity and 87%
 similarity to rat NALAD-ase L. It is most highly expressed in
 small intestine, spleen and testis. Several splice variants (see
 AA258871-77) were identified during the course of cDNA cloning
 and RT-PCR gene expression analysis. The invention provides human
 NALAD-ase L, II and IV cDNAs and encoded polypeptides, as well as
 vectors, host cells, transgenic organisms, antisense nucleic acids,
 agonists and antagonists. These are useful for treating neural
 disorders such as Alzheimer's disease, schizophrenia, ALS,
 Parkinson's disease, peripheral neuropathy, Huntington's disease,
 acute brain injury, multiple sclerosis, exposure to neurotoxins,
 peripheral nerve trauma, ischemia or dementia (claimed).

Sequence 740 AA;

Query Match 75.6%; Score 34; DB 21; Length 740;
 Best Local Similarity 71.4%; Pred. No. 2.9e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 FSAMGAE 8
 1:|||||
 Db 410 faswage 416

RESULT 14
 AAB12947

ID AAB12947 standard; peptide; 8 AA.

XX AAB12947;

DT 23-NOV-2000 (first entry)

DE Helicokinin hormone (HK-III) peptide.

KW Crop pest; helicokinin; hormone; corn earworm; growth retardation;
 HK-III.

OS Synthetic.

PN US6087165-A.

XX 11-JUL-2000.

PD 26-FEB-1999; 99US-0258275.

PF 26-FEB-1999; 99US-0258275.

PR (USDA) US SEC OF AGRIC.

XX Raina AK, Leclerc RF, Vakharia VN;

DR WPI: 2000-505106/45.

XX New recombinant baculovirus comprising DNA encoding a peptide hormone
 with helicokinin activity, useful for controlling Helicoverpa zea and
 other related species

PT Disclosure; Column 3; 8pp; English.

PS This invention relates to a recombinant baculovirus comprising DNA
 encoding at least one peptide hormone having helicokinin (HK) activity.
 The invention includes peptide sequence for three synthetic HK hormones
 (AAB12945-B12947) these peptides help regulate the physiological
 processes enabling the growth of Helicoverpa zea (corn earworm) larvae.
 Infestations of H. zea are a major cause of crop loss. The recombinant
 baculovirus of the invention containing DNA encoding a hormone with HK
 activity can be used for controlling populations of H. zea and other
 related crop pest species. The present sequence represents a synthetic
 helicokinin III hormone peptide. The peptide is used in the production
 of the baculovirus of the invention.

CC Sequence 8 AA;

XX

SO

Query Match 73.3%; Score 33; DB 21; Length 8;
 Best Local Similarity 83.3%; Pred. No. 3.4e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RFSAMG 6
 :|||||
 Db 3 kfsawg 8

RESULT 15
 AAR71029

ID AAR71029 standard; Peptide; 16 AA.

XX AAR71029;

AC 02-OCT-1995 (first entry)

DT HSV-2 protease substrate.

DE

XX Herpes simplex virus; HSV-2; capsid; UL26 gene.
KW Synthetic.
OS
XX
XX W09506055-A.
PN
XX
XX 02-MAR-1995.
PD
XX
XX 19-AUG-1994; 94WO-US09303.
PF
XX
XX 20-AUG-1993; 93US-0110522.
PR
XX 23-JUN-1994; 94US-0264537.
XX
XX (SMIR) SMITHKLINE BECHAM CORP.
PA
XX
XX Debouck CM, Dilella AG;
PI
XX
XX WPI; 1995-106803/14.
CR
XX
XX New herpes simplex virus (HSV)-2 protease and capsid protein
PT used to develop prods. for use in the diagnosis and treatment of
HSV-2 infections
PS
XX
XX Claim 19; Page 12; 51pp; English.
PS
XX
XX The sequence is that of a synthetic peptide substrate for the herpes
XX simplex virus type 2 gene UL26 product, the HSV-2 protease. The
XX peptide may be used to identify compounds which modulate HSV-2 protease
XX activity
XX See also AAR71017-31.
XX
XX
XX Sequence 16 AA;
XX

Query Match 73.3%; Score 33; DB 16; Length 16;
Best Local Similarity 62.5%; Pred. No. 8.6;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Y 1 RESAWGAE 8
: | | | | |
b 9 ktkmgae 16

Search completed: October 18, 2001, 16:51:16
Job time: 41 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2001, 16:50:35 ; Search time 18.9 seconds
(without alignments)
8.715 Million cell updates/sec

Title: US-09-646-579-1

Perfect score: 45

Sequence: 1 RFSANGAE 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

num DB seq length: 0
num DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*\n2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*\n3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*\n4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*\n5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep:*\n6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	73.3	8	US-09-258-275-5	Sequence 5, Appl1
2	33	73.3	19	US-09-167-434-15	Sequence 15, Appl1
3	33	73.3	19	US-08-853-755-15	Sequence 15, Appl1
4	33	73.3	20	US-07-798-776-12	Sequence 12, Appl1
5	33	73.3	20	US-08-251-288A-12	Sequence 12, Appl1
6	33	73.3	20	US-09-298-819A-12	Sequence 12, Appl1
7	33	73.3	306	US-08-279-754-2	Sequence 2, Appl1
8	33	73.3	306	PCT-US95-09052-2	Sequence 2, Appl1
9	33	73.3	350	US-08-458-067-2	Sequence 2, Appl1
10	33	73.3	350	PCT-US96-07795-2	Sequence 2, Appl1
11	33	73.3	635	PCT-US96-07796-2	Sequence 2, Appl1
12	33	73.3	635	US-07-832-855-2	Sequence 2, Appl1
13	33	73.3	1150	US-08-589-756-3	Sequence 3, Appl1
14	33	73.3	1164	US-08-589-756-1	Sequence 1, Appl1
15	33	73.3	1167	US-08-589-756-2	Sequence 2, Appl1
16	33	73.3	657	US-08-508-761B-2	Sequence 2, Appl1
17	32	71.1	1261	US-08-252-966B-18	Sequence 18, Appl1
18	31	68.9	7	US-09-258-275-4	Sequence 4, Appl1
19	31	68.9	22	US-08-310-912A-81	Sequence 81, Appl1
20	31	68.9	22	US-08-841-089-81	Sequence 81, Appl1
21	31	68.9	22	PCT-US95-04570-81	Sequence 81, Appl1
22	31	68.9	22	PCT-US95-04570-81	Sequence 81, Appl1
23	31	68.9	280	US-08-409-452-1	Sequence 81, Appl1
24	30.5	67.8	10	US-08-159-339A-1098	Sequence 1098, Ap
25	30.5	67.8	12	US-08-735-253-1	Sequence 11, Appl1
26	30.5	67.8	20	US-08-468-540B-11	Sequence 23, Appl1
27	30.5	67.8	20	US-08-297-395-23	Sequence 23, Appl1

28	30.5	67.8	168	6	5194425-4	Patent No. 5194425
29	30.5	67.8	170	1	US-08-227-372-1	Sequence 1, Appl1
30	30.5	67.8	170	2	US-08-327-357A-1	Sequence 1, Appl1
31	30.5	67.8	170	3	US-08-470-397-1	Sequence 1, Appl1
32	30.5	67.8	170	3	US-09-007-520-1	Sequence 1, Appl1
33	30.5	67.8	170	3	US-08-462-351-3	Sequence 3, Appl1
34	30.5	67.8	170	3	US-09-055-263-1	Sequence 1, Appl1
35	30.5	67.8	170	6	5194425-3	Patent No. 5194425
36	30.5	67.8	170	6	5468481-3	Patent No. 5468481
37	30.5	67.8	171	2	US-08-781-122-2	Sequence 2, Appl1
38	30.5	67.8	171	4	US-09-137-759-2	Sequence 2, Appl1
39	30	66.7	118	3	US-09-012-084-4	Sequence 4, Appl1
40	30	66.7	130	3	US-08-906-769-157	Sequence 157, App
41	30	66.7	130	3	US-08-906-616-157	Sequence 157, App
42	30	66.7	130	4	US-08-639-075A-157	Sequence 157, App
43	30	66.7	130	4	US-09-012-431-157	Sequence 157, App
44	30	66.7	130	4	US-09-012-692-157	Sequence 157, App
45	30	66.7	130	4	US-08-906-613-157	Sequence 157, App

ALIGNMENTS

RESULT 1
US-09-258-275-5
Sequence 5, Application US/09258275
Patent No. 6087165
GENERAL INFORMATION:
APPLICANT: Raina, Ashok A.
APPLICANT: Leclerc, Robert F.
APPLICANT: Vakharia, Vikram N.
TITLE OF INVENTION: Recombinant Baculovirus and its use as a Biocontrol
FILE REFERENCE: 0140.97
CURRENT APPLICATION NUMBER: US/09/258,275
CURRENT FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 8
TYPE: PRT
ORGANISM: Helicoverpa zea
US-09-258-275-5

Query Match 73.3%; Score 33; DB 3; Length 8;
Best Local Similarity 83.3%; Pred. No. 1.5e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RFSANG 6
DB 3 KFSANG 8

RESULT 2
US-09-167-434-15
Sequence 15, Application US/09167434
Patent No. 6008033
GENERAL INFORMATION:
APPLICANT: Abdel-Meguid, Sherin
APPLICANT: Qiu, Xiyang
APPLICANT: Culp, Jeffrey
APPLICANT: Debouck, Christine
APPLICANT: Janson, Cheryl A.
APPLICANT: Hoop, Susan S.
APPLICANT: Smith, Ward W.
TITLE OF INVENTION: No. 6008033el Proteases, Compositions Capable of
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road - P.O. Box 1539
CITY: King of Prussia

STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/167,434
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/853,755
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/022,470
FILING DATE: 26-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,416
FILING DATE: 21-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030,901
FILING DATE: 14-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/035,973
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,191
FILING DATE: 27-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: P50472-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5022
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
S-09-167-434-15

Query Match 73.3%; Score 33; DB 3; Length 19;
Best Local Similarity 62.5%; Pred. No. 8.2;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Y 1 RESAWGAE 8
: | ||||
b 3 KEKINGAE 10

ESULT 3
S-08-853-755-15
Sequence 15, Application US/08853755
Patent No. 6083711
GENERAL INFORMATION:
APPLICANT: Abdel-Meguid, Sherin
APPLICANT: Qiu, Xinyang
APPLICANT: Culp, Jeffrey
APPLICANT: Debouck, Christine
APPLICANT: Janson, Cheryl A.
APPLICANT: Hoog, Susan S.
APPLICANT: Smith, Ward W.
TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of
TITLE OF INVENTION: Binding to Said Site, and Methods of Use thereof
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smltkline Beecham Corporation

STREET: 709 Swedeland Road - P.O. Box 1539
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,755
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/018,616
FILING DATE: 15-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/022,470
FILING DATE: 26-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,416
FILING DATE: 21-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030,901
FILING DATE: 14-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/035,973
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,191
FILING DATE: 27-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: P50472-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5022
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-853-755-15

Query Match 73.3%; Score 33; DB 3; Length 19;
Best Local Similarity 62.5%; Pred. No. 8.2;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RESAWGAE 8
: | ||||
DB 3 KEKINGAE 10

RESULT 4
US-07-798-776-12
Sequence 12, Application US/07798776
Patent No. 5434074
GENERAL INFORMATION:
APPLICANT: GIBSON, D. WADE
APPLICANT: WELCH, ANTHONY R.
TITLE OF INVENTION: HERPES VIRUS PROTEINASE AND METHOD OF
TITLE OF INVENTION: ASSAYING
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W., Eleventh Floor
CITY: Washington
STATE: D.C.

COUNTRY: U.S.A.
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/798,776
FILING DATE: 19911127
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FOULKE, CYNTHIA L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: 1107.07080
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 508-9100
TELEFAX: (202) 508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Herpes simplex virus type 1
JS-07-798-776-12

Query Match 73.3%; Score 33; DB 1; Length 20;
Best local similarity 62.5%; Pred. No. 8.7;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 RFSAMGAE 8
13 KFKMGAE 20

RESULT 5
JS-08-251-288A-12
Sequence 12, Application US/08251288A
Patent No. 6001967
GENERAL INFORMATION:
APPLICANT: Gibson, Wade
TITLE OF INVENTION: HERPES PROTEINASE AND METHOD
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Wilcoff
STREET: 1001 G Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/251,288A
FILING DATE: 31-MAY-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 01107.46284
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6001967e
US-08-251-288A-12

Query Match 73.3%; Score 33; DB 3; Length 20;
Best local similarity 62.5%; Pred. No. 8.7;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 RFSAMGAE 8
13 KFKMGAE 20

RESULT 6
US-09-298-819A-12
Sequence 12, Application US/09298819A
Patent No. 6077679
GENERAL INFORMATION:
APPLICANT: Gibson, Wade
TITLE OF INVENTION: HERPES PROTEINASE AND METHOD
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Wilcoff
STREET: 1001 G Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/298,819A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/251,288
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 01107.46284
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6077679e
US-09-298-819A-12

Query Match 73.3%; Score 33; DB 3; Length 20;
Best Local Similarity 62.5%; Pred. No. 8.7;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Y 1 RFSAWGAE 8
: 1 1111
b 13 KFKMGAE 20

ESULT 7

3-08-279-754-2
Sequence 2, Application US/08279754
Patent No. 5486470

GENERAL INFORMATION:

APPLICANT: DARKE, PAUL L.

APPLICANT: HALL, DAMN L.

APPLICANT: KUO, LAWRENCE C.

TITLE OF INVENTION: PURIFIED HERPES SIMPLEX VIRAL PROTEASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESS: P.O. BOX 2000, 126 E. LINCOLN AVE.

CITY: RAHWAY

STATE: NJ

COUNTRY: USA

ZIP: 07065

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/279,754

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: GIESSEER, JOANNE M.

REGISTRATION NUMBER: 32,838

REFERENCE/DOCKET NUMBER: 19233

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3046

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 306 amino acids

TYPE: amino acid

STRANDEDNESS: double

TOPOLOGY: linear

HYPOTHETICAL: NO

ANTI-SENSE: NO

3-08-279-754-2

Query Match 73.3%; Score 33; DB 1; Length 306;

Best Local Similarity 62.5%; Pred. No. 1.2e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 RFSAWGAE 8

: 1 1111

250 KFKMGAE 257

Sequence 2, Application PC/US9509052

GENERAL INFORMATION:

APPLICANT: DARKE, PAUL L.

APPLICANT: HALL, DAMN L.

APPLICANT: KUO, LAWRENCE C.

TITLE OF INVENTION: PURIFIED HERPES SIMPLEX VIRAL PROTEASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESS: JOANNE M. GIESSEER

STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.

CITY: RAHWAY

STATE: NJ

COUNTRY: USA

ZIP: 07065

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/09052

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: GIESSEER, JOANNE M.

REGISTRATION NUMBER: 32,838

REFERENCE/DOCKET NUMBER: 19233

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3046

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 306 amino acids

TYPE: amino acid

STRANDEDNESS: double

TOPOLOGY: linear

HYPOTHETICAL: NO

ANTI-SENSE: NO

PCT-US95-09052-2

Query Match 73.3%; Score 33; DB 5; Length 306;

Best Local Similarity 62.5%; Pred. No. 1.2e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 RFSAWGAE 8

: 1 1111

250 KFKMGAE 257

Sequence 2, Application US/08458067

Patent No. 5728557

GENERAL INFORMATION:

APPLICANT: Register, Robert B.

APPLICANT: Shafer, Jules A.

TITLE OF INVENTION: HERPES SIMPLEX TYPE 1 PROTEASE MUTANTS

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESS: Ms. Joanne M. Giesseer

STREET: 126 East Lincoln Avenue, P.O. Box 2000-0907

CITY: Rahway

STATE: New Jersey

COUNTRY: US

ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/458,067

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Giesseer, Joanne M.

REGISTRATION NUMBER: 32,838

REFERENCE/DOCKET NUMBER: 19457

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3046

TELEFAX: (908) 594-4720
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 350 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-458-067-2

Query Match 73.3%; Score 33; DB 1; Length 350;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RESAMGAE 8
: 1 1111
Db 250 KFKMGAE 257

RESULT 10

PCT-US96-07795-2
Sequence 2, Application PC/TUS9607795
GENERAL INFORMATION:

APPLICANT: MERCK & CO., INC.
APPLICANT: Register, Robert B.
APPLICANT: Shafer, Jules A.
TITLE OF INVENTION: HERPES SIMPLEX TYPE 1 PROTEASE MUTANTS
TITLE OF INVENTION: AND VECTORS
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:

ADDRESSEE: Ms. Joanne M. Giesser
STREET: 126 East Lincoln Avenue, P.O. Box 2000-0907
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/07795

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Giesser, Joanne M.

REGISTRATION NUMBER: 32,838

REFERENCE/DOCKET NUMBER: 19457

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3046

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 350 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-07795-2

Query Match 73.3%; Score 33; DB 5; Length 350;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RESAMGAE 8
: 1 1111
Db 250 KFKMGAE 257

RESULT 11
PCT-US96-07796-2
Sequence 2, Application PC/TUS9607796

GENERAL INFORMATION:
APPLICANT: MERCK & CO., INC.
APPLICANT: Register, Robert B.
APPLICANT: Shafer, Jules A.
TITLE OF INVENTION: HERPES SIMPLEX TYPE 1 PROTEASE MUTANTS
TITLE OF INVENTION: AND VECTORS
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:

ADDRESSEE: Ms. Joanne M. Giesser
STREET: 126 East Lincoln Avenue, P.O. Box 2000-0907
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/07796

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Giesser, Joanne M.

REGISTRATION NUMBER: 32,838

REFERENCE/DOCKET NUMBER: 19457

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3046

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 350 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-07796-2

Query Match 73.3%; Score 33; DB 5; Length 350;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RESAMGAE 8
: 1 1111
Db 250 KFKMGAE 257

RESULT 12

US-07-832-855-2

Sequence 2, Application US/07832855

Patent No. 5478727

GENERAL INFORMATION:

APPLICANT: Rolzman, Bernard

APPLICANT: Liu, Fenyong

TITLE OF INVENTION: Methods and Compositions of a

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE & DURKEE

STREET: 321 No. 5478727th Clark Street, Suite 800

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60610

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/832.855

FILING DATE: 19920207

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cooley, Ronald B.
REGISTRATION NUMBER: 27,187
REFERENCE/DOCKET NUMBER: ARCD045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 744-0090
TELEFAX: (312) 245-4961
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
S-07-832-855-2

Query Match
st Local Similarity 73.3%; Score 33; DB 1; Length 635;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Y 1 RFSWAG 8
250 RFSWAG 257

RESULT 13
S-08-589-756-3
Sequence 3, Application US/08589756
Patent No. 5846547
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: STREPTOCOCCAL C5a PEPTIDASE VACCINE
NUMBER OF SEQUENCES: 5
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/589,756
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
S-08-589-756-3

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 1150;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 RFSWAG 6
473 RFSWAG 478

RESULT 14
S-08-589-756-1
Sequence 1, Application US/08589756
Patent No. 5846547
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: STREPTOCOCCAL C5a PEPTIDASE VACCINE
NUMBER OF SEQUENCES: 5
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/589,756
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1164 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-589-756-1

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 1164;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 RFSWAG 6
473 RFSWAG 478

RESULT 15
US-08-589-756-2
Sequence 2, Application US/08589756
Patent No. 5846547
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: STREPTOCOCCAL C5a PEPTIDASE VACCINE
NUMBER OF SEQUENCES: 5
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/589,756
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1167 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-589-756-2

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 1167;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 RFSWAG 6
473 RFSWAG 478

Search completed: October 18, 2001, 16:51:41
Job time: 66 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2001, 16:50:35 ; Search time 22.89 Seconds
(without alignments)
26.623 Million cell updates/sec

Title: US-09-646-579-1

Perfect score: 45

Sequence: 1 F8SAGAE 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

num DB seq length: 0
num DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR68:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	82.2	553	2	D96684
2	36	80.0	500	2	H70629
3	35	77.8	284	2	S66427
4	35	77.8	536	2	B83278
5	34	75.6	172	2	PN0345
6	34	75.6	357	2	E75284
7	34	75.6	545	2	D75208
8	33	73.3	272	2	D83154
9	33	73.3	306	2	A54862
10	33	73.3	386	2	T47224
11	33	73.3	576	1	S22453
12	33	73.3	635	1	WMBE6
13	33	73.3	637	2	A56263
14	33	73.3	1167	1	A35066
15	32	71.1	171	2	S09903
16	32	71.1	171	2	S09759
17	32	71.1	173	2	H82141
18	32	71.1	254	2	S48497
19	32	71.1	272	2	T24709
20	32	71.1	311	2	B69191
21	32	71.1	383	2	H84700
22	32	71.1	405	2	A85485
23	32	71.1	405	2	S40559
24	32	71.1	431	2	A72037
25	32	71.1	431	2	D86588
26	32	71.1	481	2	E82372
27	32	71.1	485	2	C75460
28	32	71.1	656	2	T18542
29	32	71.1	657	2	S25184

30	32	71.1	857	2	S16133	dimethylglycine de
31	32	71.1	1229	2	A56068	co-repressor prote
32	31	68.9	132	2	PN0546	gamma5-crystallin
33	31	68.9	135	1	ASL2W	nef protein (clone
34	31	68.9	165	1	ASL22	nef protein - equi
35	31	68.9	165	2	C45711	reverse transcript
36	31	68.9	168	1	CYFG2	gamma-crystallin I
37	31	68.9	169	2	T01783	gamma-2-crystallin
38	31	68.9	175	2	JN0682	gamma2-crystallin
39	31	68.9	175	2	JN0681	gamma2-crystallin
40	31	68.9	176	2	S45015	gamma-crystallin M
41	31	68.9	257	2	D82751	conserved hypothet
42	31	68.9	267	2	S70643	dolichyl-phosphate
43	31	68.9	296	2	T32222	hypothetical prote
44	31	68.9	348	2	C83571	probable binding p
45	31	68.9	352	2	F83617	hypothetical prote

ALIGNMENTS

RESULT 1
D96684
hypothetical protein F12P19.16 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96684
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzla, R.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: D96684
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-553 <STO>
A:Cross-references: GB:AEO05173; NID:96227024; PIDN:AF06060.1; GSPDB:GN00141
C:Genetics:
A:Gene: F12P19.16
A:Map position: 1

Query Match
Best Local Similarity 82.2% Score 37; DB 2; Length 553;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 F8SAGAE 8
Db 407 F8SAGAE 413

RESULT 2
H70629
probable AMINOPEPTIDASE - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: H70629
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Fieldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, R.; Rastaud, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:9825987
A:Accession: H70629
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

:Residues: 1-500 <COL>
 :Cross-references: GB:284724; GB:AL23456; NID:93261708; PIDN:CAB06558.1; PID:g1817684
 :Experimental source: strain H37Rv
 :Genetics:
 :Gene: lpgL

Query Match 80.0%; Score 36; DB 2; Length 500;
 Best Local Similarity 75.0%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 RFSAMGAE 8
 ||:||||
 309 RFAFWGAE 316

RESULT 3

:peptidase (EC 3.4.11.-) - Streptomyces griseus
 :Species: Streptomyces griseus

Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 31-Mar-2000

Accession: S66427
 :Matas, B.; Greenblatt, H.M.; Shoham, G.; Spungin-Bialik, A.; Blumberg, S.; Barra, D.
 :J. Biochem. 236, 843-846, 1996

Title: Aminopeptidase from Streptomyces griseus. Primary structure and comparison with
 :Reference number: S66427; MUID:96270734

:Accession: S66427
 :Status: preliminary
 :Molecule type: protein
 :Residues: 1-284 <MAR>
 :Keywords: alpha-aminoacylpeptide hydrolase; zinc; zymogen

Query Match 77.8%; Score 35; DB 2; Length 284;
 Best Local Similarity 75.0%; Pred. No. 52;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 RFSAMGAE 8
 ||:||||
 124 RFAFWGAE 131

RESULT 4
 33278
 :able aminopeptidase PA2939 [imported] - Pseudomonas aeruginosa (strain PA01)

:Accession: S63278
 :Stover, C.K.; Plam, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
 :Lory, S.; Olson, M.V.
 :J. Biol. Chem. 269, 959-964, 1994

Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 :Reference number: A82950; MUID:20437337
 :Accession: S63278
 :Status: preliminary
 :Molecule type: DNA
 :Residues: 1-536 <STO>
 :Cross-references: GB:AE004720; GB:AE004091; NID:99949032; PIDN:AG06327.1; GSPDB:GN001
 :Experimental source: strain PA01
 :Genetics:
 :Gene: PA2939

Query Match 77.8%; Score 35; DB 2; Length 536;
 Best Local Similarity 75.0%; Pred. No. 94;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 RFSAMGAE 8
 ||:||||
 333 RFAFWGAE 340

RESULT 5
 PN0545
 gamma4-crystallin - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Nov-1997

C:Accession: PN0545

R:Smolich, B.D.; Tarkington, S.K.; Saha, M.S.; Stachakis, D.G.; Gralinger, R.M.
 Gene 128, 189-195, 1993

A>Title: Characterization of Xenopus laevis gamma-crystallin-encoding genes.

A:Reference number: JN0680; MUID:93292986

A:Accession: PN0545

A:Molecule type: DNA

A:Residues: 1-172 <SMO>

A>Note: the authors translated the codon CAT for residue 82 as Gln and CAC for residu

C:Gene: gamma4 cry
 A:introns: 37/3; 85/3; 126/3
 C:Superfamily: beta-crystallin
 C:Keywords: duplication

Query Match 75.6%; Score 34; DB 2; Length 172;
 Best Local Similarity 85.7%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 RFSAMGA 7
 ||:||||
 Db 151 RFSMDGA 157

RESULT 6

E75284
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: E75284

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999

A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: E75284

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-357 <WHI>
 A:Cross-references: GB:AE002066; GB:AE000513; NID:96460163; PIDN:AAF11904.1; PID:9646

A:Experimental source: strain R1

C:Genetics:
 A:Gene: DR2359
 A:Map position: 1

Query Match 75.6%; Score 34; DB 2; Length 357;
 Best Local Similarity 62.5%; Pred. No. 96;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 RFSAMGAE 8
 ||:||||
 Db 123 RFAFWGAE 130

RESULT 7

D75208
 sugar abc transporter binding protein related PAB2439 - Pyrococcus abyssi (strain Ots

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: D75208

R:anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome s

A:Reference number: A75001

A:Accession: D75208

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-545 <RAW>
A:Cross-references: GB:AJ248283; GB:AL096836; NID:95457433; PIDN:CAB49115.1; PID:9545762
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB2439

Query Match
Best Local Similarity 75.6%; Score 34; DB 2; Length 545;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 RFSAMGA 7
DB 63 KFSAMGS 69

RESULT 8

DB3154
Probable permease of ABC taurine transporter PA3936 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Genetics:
A:Gene: C.K.; Pham, X.Q.; Ervin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lam,
L.; Lory, S.; Olson, M.V.
Nucleotide 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337

A:Accession: DB3154

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-272 <STG>

A:Cross-references: GB:AE004811; GB:AE004091; NID:99950116; PIDN:AG07323.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA3936

C:Superfamily: Synecchococcus nitrate transport protein nrtB

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 272;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

2Y 1 RFSAMG 6
DB 267 RFSAMG 272

RESULT 9

tsv-1 proteinase - human herpesvirus 1
A:Species: human herpesvirus 1
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 16-Feb-1997
A:Accession: AS4862

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337

A:Accession: AS4862

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-306 <DAR>

A:Cross-references: GB:LJ32018

C:Superfamily: varicella-zoster virus gene 33 protein

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 306;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 RFSAMGA 8

DB 250 KFSAMGA 257

RESULT 10

O-succinylhomoserine (thiol)-lyase (EC 4.2.99.9) [validated] - Corynebacterium glutam
C:Species: Corynebacterium glutamicum

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000

C:Accession: T47234

A:Title: Analysis of corynebacterium glutamicum methionine biosynthetic pathway: Isol

A:Reference number: Z24410; MUID:99347339

A:Accession: T47234

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-386 <HMA>

A:Cross-references: EMBL:AF126953; NID:95852461; PIDN:AD54070.1; PID:95852462

A:Experimental source: strain AS019

C:Genetics:

A:Gene: metB

C:Function:

A:Description: EC 4.2.99.9 [validated; MUID:99347339]

A:Pathway: methionine biosynthesis

A:Note: cofactor pyridoxal-phosphate

C:Superfamily: O-succinylhomoserine (thiol)-lyase

C:Keywords: carbon-oxygen lyase

Query Match
Best Local Similarity 71.4%; Score 33; DB 2; Length 386;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

2Y 2 RFSAMGA 8
DB 118 RFSAMGA 124

RESULT 11

colicin E7 (EC 3.1.21.-) - Escherichia coli plasmid COLE7-K317
C:Species: Escherichia coli
C:Date: 04-Dec-1992 #sequence_revision 24-Apr-1998 #text_change 11-Jun-1999

C:Accession: S22453; S49179

R:Soong, B.W.; Lu, F.M.; Chak, K.F.

Mol. Gen. Genet. 233, 177-183, 1992

A:Title: Characterization of the cea gene of the COLE7 plasmid.

A:Reference number: S22453; MUID:92293113

A:Accession: S22453

A:Molecule type: DNA

A:Residues: 1-576 <SOO>

A:Cross-references: EMBL:M62409; NID:9144374; PIDN:AAA8054.1; PID:9144375

A:Note: the authors translated the codon GAA for residue 366 as Phe

R:lan, P.C.K.; Parsons, M.

submitted to the EMBL Data Library, December 1991

A:Description: Nucleotide sequence encoding the immunity and lysis proteins and the c

A:Reference number: S49176

A:Accession: S49179

A:Molecule type: DNA

A:Residues: 371-576 <LAD>

A:Cross-references: EMBL:X63620; NID:9510384; PIDN:CAA45164.1; PID:9510385

C:Genetics:

A:Gene: cea

A:Superfamily: Plasmid COLE7-K317

C:Keywords: bacteriocin; endonuclease; hydrolase; toxin

Query Match
Best Local Similarity 73.3%; Score 33; DB 1; Length 576;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 RFSAMG 6
:|||||
b 123 KFSAMG 128

ESULT 12
MBEW6
apsid protein - human herpesvirus 1 (strain 17)
Species: human herpesvirus 1
Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
Accession: H30084
McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perr
Gen. Virol. 69, 1531-1574, 1988
Title: The complete DNA sequence of the long unique region in the genome of herpes sim
Reference number: A30083; MUID:88274327
Accession: H30084
tus: nucleic acid sequence not shown; translation not shown
ecule type: DNA
Residues: 1-635 <MCG>
Cross-references: GB:X14112; NID:91944536; PIDN:CAA32318.1; PID:959526; GB:D00317
Genetics:
Gene: UL26
Superfamily: varicella-zoster virus gene 33 protein
Keywords: capsid protein

Query Match 73.3%; Score 33; DB 1; Length 635;
Best Local Similarity 62.5%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Y 1 RFSAMGAE 8
:|||||
b 250 KFKMGAE 257

ESULT 13
56263
eta-galactosidase (EC 3.2.1.23) isozyme 12 - Arthrobacter sp. (strain B7)
Species: Arthrobacter sp.
Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 17-Mar-2000
Accession: A56263
Guthall, K.R.; Trimbur, D.E.; Kasmit, J.J.; Branchley, J.E.
Bacteriol. 177, 1981-1988, 1995
Title: Analysis of a novel gene and beta-galactosidase isozyme from a psychrotrophic A
Reference number: A56263; MUID:95338267
Accession: A56263
Status: preliminary
Molecule type: DNA
Residues: 1-637 <GUM>
Cross-references: GB:U17417; NID:9676845; PIDN:AAA75601.1; PID:9676846
Superfamily: Bacillus beta-galactosidase
Keywords: glycosidase; hydrolase

Query Match 73.3%; Score 33; DB 2; Length 637;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 RFSAMG 6
:|||||
b 245 RFSWG 250

ESULT 14
35066
Streptococcal C5a peptidase (EC 3.4.21.-) precursor - Streptococcus pyogenes
Species: Streptococcus pyogenes
Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 18-Jun-1999
Accession: A35066; S52539
Chen, C.C.; Cleary, P.P.
Biol. Chem. 265, 3161-3167, 1990
Title: Complete nucleotide sequence of the streptococcal C5a peptidase gene of Streptococcus

A:Reference number: A35066; MUID:90153964
A:Accession: A35066
A:Molecule type: DNA
A:Residues: 1-1167 <CHE>
A:Cross-references: GB:J05229; NID:9153776; PIDN:AAA26960.1; PID:9552009
A>Note: the amino end of the mature protein was determined by protein sequencing
R:Katerov, V.; Schalen, C.; Totolian, A.A.
Mol. Gen. Genet. 245, 78-85, 1994
A:Title: Sequencing of genes within the vir regulon of Streptococcus pyogenes type M1
A:Reference number: S52535; MUID:95147851
A:Accession: S52539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75, 'V', '77-78, 'K', '80-83, 'PS', <KAT>
A:Cross-references: GB:S75411; NID:9914107; PIDN:AAB33264.1; PIV:9914112
A>Note: in Genbank entry S75411, release 106.0, the initiation codon TTG for residue
C:Genetics:
A:Gene: scpA
A:Start codon: TTG
C:Function:
A:Description: specifically cleaves a Lys-Asp peptide bond near the carboxyl end of h
C:Superfamily: streptococcal C5a peptidase; subtilisin homology
C:Keywords: cell wall; hydrolase; serine proteinase; tandem repeat; transmembrane pro
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-1167/Product: streptococcal C5a peptidase #status predicted <MAT>
F:121-526/Domain: subtilisin homology #status atypical <SBT>
F:1029-1139/Domain: cell wall spanning #status predicted <CMS>
F:1034-1101/Region: glycine/proline-rich 17-residue repeats
F:1040-1157/Domain: transmembrane #status predicted <TMM>
F:130,193,512/Active site: Asp, His, Ser #status predicted

Query Match 73.3%; Score 33; DB 1; Length 1167;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 RFSAMG 6
:|||||
b 473 RFSWG 478

ESULT 15
S09903
hypothetical protein IRL10 precursor - human cytomegalovirus (strain AD169)
C:Species: human cytomegalovirus, human herpesvirus 5
A>Note: host Homo sapiens (man)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C:Accession: S09903
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T
M.; Barrett, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovi
A:Reference number: S09749; MUID:90269039
A:Accession: S09903
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-171 <CHE>
A:Cross-references: EMBL:X17403; NID:959591; PIDN:CAA35300.1; PID:91780920
A>Note: this sequence was submitted to the EMBL Data Library, December 1989
C:Keywords: glycoprotein; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-171/Product: hypothetical protein IRL10 #status predicted <MAT>
F:70-103/Domain: transmembrane #status predicted <TMM>
F:48,49,56,108/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.1%; Score 32; DB 2; Length 171;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 2 RFSAMG 7
:|||||
b 78 RFSAMG 83

Thu Oct 18 16:57:42 2001

Search completed: October 18, 2001, 16:52:11
Job time: 96 sec

us-09-646-579-1.rpt

Page 5

T METAL 132 132 ZINC 2.
T METAL 160 160 ZINC 1.
T METAL 247 247 ZINC 2.
T DISULFID 245 250
Q SEQUENCE 284 AA; 29722 MW; 67F1B80F8CA5C4CC CRC64;

Query Match 77.8%; Score 35; DB 1; Length 284;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 RESANGAE 8
11: 1111
b 124 REAWGAE 131

RESULT 2
RCG4_XENLA STANDARD; PRT; 172 AA.
P55941;

T 01-NOV-1997 (Rel. 35, Created)
T 01-NOV-1997 (Rel. 35, Last sequence update)
T 01-NOV-1997 (Rel. 35, Last annotation update)
E GAMMA CRYSTALLIN IV (GAMMA-4-CRY) (FRAGMENT).
N GCRY4.

S Xenopus laevis (African clawed frog).
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
C Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
C Xenopodinae; Xenopus.
X NCBI_TaxID=8355;
[1]

SEQUENCE FROM N.A.

X MEDLINE-93292986; PubMed-8514186;
X Smolich B.D., Tarkington S.K., Saha M.S., Stathakis D.G.,
A Granger R.M.;

Characterization of Xenopus laevis gamma-crystallin-encoding genes.;
L Gene 128:189-195(1993).
T 1- FUNCTION: CRYSTALLINS ARE THE DOMINANT STRUCTURAL COMPONENTS
C OF THE VERTEBRATE EYE LENS.

-1- SUBUNIT: MONOMER (BY SIMILARITY).
C -1- DOMAIN: HAS A TWO-DOMAIN BETA-STRUCTURE, FOLDED INTO FOUR
C VERY SIMILAR GREEK KEY MOTIFS.

-1- SIMILARITY: BELONGS TO THE BETA/GAMMA-CRYSTALLIN FAMILY.
C -----

C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
C or send an email to license@sib-sib.ch).

C EMBL; M9582; -; NOT_ANNOTATED_CDS.
R HSSP; P02526; 1GCS.

R InterPro; IPR001064; -;
R Pfam; PF00030; crystal1.2.

R PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 4.
W Eye lens protein; Duplication; Multigene family.
T NON_TER 1
T DOMAIN <1 37 MOTIF 1.
T DOMAIN 38 81 MOTIF 2.
T DOMAIN 82 85 CONNECTING PEPTIDE.
T DOMAIN 86 126 MOTIF 3.
T DOMAIN 127 172 MOTIF 4.

Q SEQUENCE 172 AA; 21274 MW; 1A0BBD094F9A6AE CRC64;

Query Match 75.6%; Score 34; DB 1; Length 172;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 RESANGA 7
111 111

DB 151 RESDWGA 157

RESULT 3
CEA7_ECOLI

ID CEA7_ECOLI STANDARD; PRT; 576 AA.
AC 047112; Q51604;
DR 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE COLICIN E7 (EC 3.1.21.1).

GN COLE7 OR CEA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.

OC Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-K317;
RC MEDLINE-92293113; PubMed-1603061;
RX Soong B.W., Lu F.M., Chak K.-F.;

RT *Characterization of the cea gene of the Cole7 plasmid.*;
RL Mol. Gen. Genet. 233:177-183(1992).
RN [2]
RP SEQUENCE OF 371-576 FROM N.A.

RC STRAIN-K317;
RA Lau P.C.K., Parsons M.;

RT *Nucleotide sequence encoding the immunity and lysis proteins and
RT the carboxyl-terminal peptides of colicins E4 and E7.*;
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
RN [3]

RX X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 371-576;
RA MEDLINE-99148115; PubMed-10368275;
KO T.P., Liao C.C., Ku W.-Y., Chak K.-F., Yuan H.S.;

RT *The crystal structure of the DNase domain of colicin E7 in complex
RT with its inhibitor Im7 protein.*;
RL Structure 7:91-102(1999).

-1- FUNCTION: THIS PLASMID-CODED BACTERICIDAL PROTEIN IS AN
CC ENDONUCLEASE ACTIVE ON BOTH SINGLE- AND DOUBLE-STRANDED DNA BUT
CC WITH UNDEFINED SPECIFICITY.

-1- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE
CC AGAINST, ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.
CC -1- SIMILARITY: BELONGS TO THE NUCLEASE FAMILY OF COLICINS AND
CC PYOSINS.

C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
C or send an email to license@sib-sib.ch).

CC EMBL; M62409; AAA98054.1; -;
DR EMBL; X63620; CAA45164.1; -;
DR PDB; 7CEI; 17-SEP-99.

DR InterPro; IPR002711; -;
DR InterPro; IPR003058; -;
DR Pfam; PF01844; HNH; 1.

DR PRINTS; PR01295; CLOACIN.
KW Antibiotic; Bacteriocin; Plasmid; Hydrolase; Endonuclease;
KW Zinc; Metal-binding; 3D-structure.

FT METAL 544 ZINC.
FT METAL 569 ZINC.
FT METAL 573 ZINC.

Q SEQUENCE 576 AA; 61349 MW; E5B05E73B2E17249 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 576;
Best Local Similarity 83.3%; Pred. No. 75;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

T DOMAIN 1127 1131 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
T SEQUENCE 1167 AA; 128263 MW; D2DDC52E5752DA5D CRC64;
Q

Query Match 73.3%; Score 33; DB 1; Length 1167;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 RFSAMG 6
111:11
b 473 RFSWAG 478

ESULT 6
R10.HCMVA STANDARD; PRT; 171 AA.
C P16808;
01-AUG-1990 (Rel. 15, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
T HYPOTHETICAL PROTEIN IRL10 PRECURSOR (TRL10).
S Human cytomegalovirus (strain AD169).
C Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
C Betaherpesvirinae; Cytomegalovirus.
N NCBI_TaxID=10360;
[1]
P SEQUENCE FROM N.A.
X MEDLINE=90269039; PubMed=2161319;
A Chee M.S., Bankier A.T., Beck S., Bohnl R., Brown C.M., Cerny R.,
Horzanel T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
A Predd E., Satchell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
T Analysis of the protein-coding content of the sequence of human
T cytomegalovirus strain AD169.
L Curr. Top. Microbiol. Immunol. 154:125-169(1990).
C
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
C or send an email to license@isb-sib.ch).
C
C EMBL: X17403; CAA35300.1;
P PIR: S09903; S09903.
R PIR: S09759; S09759.
P Hypothetical protein; Glycoprotein; Transmembrane; Signal.
T SIGNAL 1 25
T CHAIN 26 171 HYPOTHETICAL PROTEIN IRL10/TRL10.
T TRANSMEM 80 100 POTENTIAL.
T CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
T CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
T CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
T CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
Q SEQUENCE 171 AA; 19034 MW; 2C2E8AD869419B86 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 171;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 2 FSAWGA 7
111111
b 78 YSAWGA 83

ESULT 7
DSB.VIBCH STANDARD; PRT; 173 AA.
C 09K006;
T 01-OCT-2000 (Rel. 40, Created)
T 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DISULFIDE BOND FORMATION PROTEIN B (DISULFIDE OXIDOREDUCTASE).
GN DSB OR VC1902.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- FUNCTION: REQUIRED FOR DISULFIDE BOND FORMATION IN SOME
CC PERIPLASMIC PROTEINS. ACTS BY OXIDIZING THE DSB PROTEIN (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE DSB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE004266; AAF95050.1;
DR TIGR; VC1902;
DR Oxidoreductase; Redox-active center; Electron transport; Chaperone;
KW Transmembrane; Inner membrane.
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 35 POTENTIAL.
FT DOMAIN 36 50 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 51 71 POTENTIAL.
FT DOMAIN 72 98 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 99 119 POTENTIAL.
FT DOMAIN 120 147 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 148 168 POTENTIAL.
FT DOMAIN 169 173 CYTOPLASMIC (POTENTIAL).
FT DISULFID 43 46 REDOX-ACTIVE (BY SIMILARITY).
FT DISULFID 107 133 REDOX-ACTIVE (BY SIMILARITY).
SQ SEQUENCE 173 AA; 19334 MW; 3B159DD2CA28D53B CRC64;

Query Match 71.1%; Score 32; DB 1; Length 173;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 FSAWGA 7
111111
Db 77 FAWGA 82

RESULT 8
YIV5 YEAST STANDARD; PRT; 254 AA.
AC P40379;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL OXIDOREDUCTASE IN LYS1-HYRI INTERGENIC REGION
DE (EC 1.-.-.-).
GN YIR035C.
OS Saccharomyces cerevisiae (Baker's yeast).

CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Barrall B.G., Baddock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Iye G., Moulle S., Moulle T., Odell C., Pearson D.,
 RA Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.,
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 DR EMBL: Z38061; CAAB6195.1;
 DR PIR: S48497; S48497.
 DR HSSP: P47227; 1BDB.
 DR SGP: S0001474; YIR035C.
 DR InterPro: IPR002198;
 DR Pfam: PF00106; adh_short; 1.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 DR Hypothetical protein; Oxidoreductase.
 FT NP BIND 6 30 NAD OR NADP (BY SIMILARITY).
 FT ACT SITE 150 150 BY SIMILARITY.
 FT SEQUENCE 254 AA; 27480 MW; 910A5460B959F78 CRC64;
 SQ
 Query Match 71.1%; Score 32; DB 1; Length 254;
 Best Local Similarity 83.3%; Pred. No. 52;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 2Y 2 FSAMGA 7
 11:111
 144 FSSMGA 149
 RESULT 9
 CAIB_ECOLI STANDARD; PRT; 405 AA.
 P31572;
 01-JUL-1993 (Rel. 26, Created)
 01-FEB-1995 (Rel. 31, Last sequence update)
 15-JUL-1998 (Rel. 36, Last annotation update)
 L-CARNITINE DEHYDRATASE (EC 4.2.1.89) (L-CDHT).
 CAIB.
 3N Escherichia coli.
 3S Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 3X Escherichia
 3X NCBI_TaxID=562;
 [1]
 SEQUENCE FROM N.A. AND SEQUENCE OF 1-19.
 STRAIN-O44 K74;
 MEDLINE-94245624; PubMed-8186598;
 RA Eichler K., Schunck W.-H., Kieber H.-P., Mandrand-Berthelot M.-A.;
 "Cloning, nucleotide sequence, and expression of the Escherichia coli
 gene encoding carnitine dehydratase.";
 JT J. Bacteriol. 176:2970-2975(1994).
 [2]
 SEQUENCE FROM N.A.
 STRAIN-K12;
 MEDLINE-92334977; PubMed-1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 A Isono K., Mizobuchi K., Nakata A.;

RT "Systematic sequencing of the Escherichia coli genome: analysis of
 the 0-2.4 min region.";
 RT Nucleic Acids Res. 20:3305-3308(1992).
 RL [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RL The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1232-1244(1997).
 CC -1- FUNCTION: CATALYZES THE ANAEROBIC DEHYDRATION OF L-CARNITINE
 (R-3-HYDROXY-4-AMINOBUTYRATE) TO CROTONBETAINE.
 CC -1- CATALYTIC ACTIVITY: L-CARNITINE = 4-(TRIMETHYLAMMONIO)BUT-2-ENATE
 + H(2)O.
 CC -1- COFACTOR: AN UNKNOWN COFACTOR IS NECESSARY FOR ACTIVITY. THIS
 COFACTOR IS FOUND ONLY IN CELLS GROWN UNDER ANAEROBIC CONDITIONS
 AND IN PRESENCE OF CARNITINE.
 CC -1- PATHWAY: CARNITINE METABOLISM. CONVERSION OF CARNITINE TO
 GAMMA-BUTYROBETAINE.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- INDUCTION: BY L-CARNITINE OR CROTONBETAINE.
 CC -1- SIMILARITY: TO BLUE ACID-INDUCIBLE OPERON PROTEIN P (BAIP) FROM
 EUBACTERIUM SP. AND TO E. COLI YFDE.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 DR EMBL: X67748; CAAM7971.1;
 DR EMBL: X73904; CAAS2112.1;
 DR EMBL: D10483; BAA01314.1;
 DR EMBL: AE000114; AAC73149.1;
 DR PIR: S40559; S40559.
 DR Ecogene: EG11559; caib.
 KW lyase.
 FT VARIANT 187 187 V -> A (IN STRAIN O44 K74).
 FT VARIANT 302 302 T -> A (IN STRAIN O44 K74).
 FT SEQUENCE 405 AA; 45126 MW; 75583684B4B5A2DE CRC64;
 SQ
 Query Match 71.1%; Score 32; DB 1; Length 405;
 Best Local Similarity 71.4%; Pred. No. 81;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 2Y 2 FSAMGAE 8
 11:111
 33 FAEMGAE 39
 Db
 RESULT 10
 TOLB_CHLPN STANDARD; PRT; 431 AA.
 ID TOLB_CHLPN
 AC 092774; 09J031;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TOLB PROTEIN PRECURSOR
 GN TOLB OR CPN0782 OR CP1090.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 CC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83558;
 [1]
 RP SEQUENCE FROM N.A.
 RN STRAIN-CW629;
 RC MEDLINE-99206606; PubMed-10192388;

A Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 A Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 T Comparative genomes of Chlamydia pneumoniae and C. trachomatis.;
 L Nat. Genet. 21:385-389(1999).
 N (12)
 P SEQUENCE FROM N.A.
 C STRAIN-AR39;
 X MEDLINE-20150255; PubMed-10684935;
 A Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
 A White O., Hickey E.K., Peterson J., Umayam L.A., Uterback T.,
 A Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
 A Bowman C., Dodson R., Gwin M., Nelson W., Deboy R., Kolonay J.,
 A McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
 T Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 T pneumoniae AR39.;
 L Nucleic Acids Res. 28:1397-1406(2000).
 N (13)
 P SEQUENCE FROM N.A.
 C STRAIN-0138;
 X MEDLINE-20330349; PubMed-10871362;
 A Shital M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 A Shital T., Ishii K., Hattori M., Kubara S., Nakazawa T.;
 T Comparison of whole genome sequences of Chlamydia pneumoniae J138
 T from Japan and CHL029 from USA.;
 L Nucleic Acids Res. 28:2311-2314(2000).
 C -1- FUNCTION: INVOLVED IN THE TONB-INDEPENDENT UPTAKE OF PROTEINS
 C (BY SIMILARITY).
 C -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 C -1- SIMILARITY: BELONGS TO THE TOLB FAMILY.
 C -----
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C or send an email to license@isb-sib.ch).
 C -----
 R EMBL: AE001659; AAD18920.1; -
 R EMBL: AE002265; AAF38861.1; -
 R EMBL: AP002547; BAA98990.1; -
 R TIGR: CP1090; -
 M Transport; Protein transport; Periplasmic; Signal.
 T SIGNAL 1 20 POTENTIAL.
 T CHAIN 21 431 TOLB PROTEIN.
 C SEQUENCE 431 AA; 47761 MW; D5AD80B32E563D1B CRC64;

 Query Match 71.1%; Score 32; DB 1; Length 431;
 Best Local Similarity 71.4%; Pred. No. 86;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Y 1 RESAMGA 7
 b 415 KPSWGA 421

 RESULT 11
 HML_PARDE STANDARD; PRT; 631 AA.
 C P12293;
 T 01-OCT-1989 (Rel. 12, Created)
 T 01-OCT-1989 (Rel. 12, Last sequence update)
 T 15-DEC-1998 (Rel. 37, Last annotation update)
 E METHANOL DEHYDROGENASE SUBUNIT 1 PRECURSOR (Ec 1.1.99.8) (MDH LARGE
 E ALPHA SUBUNIT) (MEDH).
 N MOX.
 S Paracoccus denitrificans.
 N Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 C Paracoccus.
 C NCBI_TaxID=266;
 N (11)
 P SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.

RX MEDLINE-87307969; PubMed-3114231;
 RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.;
 RT Isolation and nucleotide sequence of the methanol dehydrogenase
 RT structural gene from Paracoccus denitrificans.;
 RL J. Bacteriol. 169:3969-3975(1987).
 C -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR - ALDEHYDE +
 C REDUCED ACCEPTOR.
 C -1- COFACTOR: POO.
 C -1- SUBCELLULAR LOCATION: PERIPLASMIC.
 C -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
 C ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
 C TOTAL CELL PROTEIN).
 C -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 C -----
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C or send an email to license@isb-sib.ch).
 C -----
 DR EMBL: M17339; AAA88366.1; -
 DR HSP, P38539; AAAH.
 DR InterPro: IPR001479; -
 DR InterPro: IPR002372; -
 DR Pfam: PF01011; Bacterial_POO_7.
 DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
 T SIGNAL 1 32
 T CHAIN 33 631 METHANOL DEHYDROGENASE SUBUNIT 1.
 T DISULFID 135 136 BY SIMILARITY.
 T DISULFID 418 447 BY SIMILARITY.
 T ACT_SITE 335 335 BASE (POTENTIAL).
 T SEQUENCE 631 AA; 69799 MW; 0934DC93FEC5730B CRC64;

 Query Match 71.1%; Score 32; DB 1; Length 631;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RESAMG 6
 Db 503 RESVWG 508

 RESULT 12
 CSPL_CORGL STANDARD; PRT; 657 AA.
 ID CSPL_CORGL 001377;
 AC 001377;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE PSI PROTEIN PRECURSOR.
 GN CSPL.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
 OC Corynebacterium.
 OX NCBI_TaxID=1718;
 OX (11)
 P SEQUENCE FROM N.A.
 C STRAIN-MELASSCOIA / ATCC 17965;
 RC MEDLINE-93023863; PubMed-1406274;
 RX Jolliff G., Mathieu L., Hahn V., Bayan N., Duchiron F., Renaud M.,
 RA Shechter E., Leblon G.;
 RT Cloning and nucleotide sequence of the cspl gene encoding PSI, one
 RT of the two major secreted proteins of Corynebacterium glutamicum: the
 RT deduced N-terminal region of PSI is similar to the Mycobacterium
 RT antigen 85 complex.;
 RL Mol. Microbiol. 6:2349-2362(1992).
 C -1- FUNCTION: ONE OF THE TWO MAJOR SECRETED PROTEINS.

CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: THE N-TERMINAL IS VERY SIMILAR TO THE COMPLETE
 CC SEQUENCES OF THE MYCOBACTERIAL ANTIGENS 85-A, 85-B & 85-C.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X66078; CAA46877.1; -
 CC PIR: S25184; S25184.
 CC InterPro: IPR000801; -
 CC Pfam: PF00756; Esterase; 1.
 CC Signal.
 CC FT CHAIN 1 43 PS1 PROTEIN.
 CC FT MOD_RES 44 44 PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
 CC SEQUENCE 657 AA; 70918 MW; 5164B87626AC098F CRC64;
 SQ
 Query Match
 Best Local Similarity 71.1%; Score 32; DB 1; Length 657;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 2 FSAMGAE 8
 1 1 1 1 1
 DB 609 FDSAMGAE 615
 RESULT 13
 M2GD_RAT STANDARD: PRT; 857 AA.
 ID M2GD_RAT
 AC 063342; -
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DIMETHYLGLYCINE DEHYDROGENASE, MITOCHONDRIAL PRECURSOR (EC 1.5.99.2)
 CE (MEZGLYDH).
 CS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
 CC NCBI_TaxID=10116;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=MISTAR;
 CC MEDLINE=91266966; PubMed=1710985;
 CC Lang H., Polster M., Brandsch R.;
 CC "Rat liver dimethylglycine dehydrogenase. Flavinylation of the enzyme
 CC in hepatocytes in primary culture and characterization of a cDNA
 CC clone.";
 CC Eur. J. Biochem. 198; 793-799(1991).
 CC -1- CATALYTIC ACTIVITY: N,N-DIMETHYLGLYCINE + ACCEPTOR + H(2)O =
 CC SARCOSINE + FORMALDEHYDE + REDUCED ACCEPTOR.
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: CHOLINE METABOLISM.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 CC -1- SIMILARITY: BELONGS TO THE GCVT FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X55995; CAA39468.1; -
 CC InterPro: IPR002536; -
 CC Pfam: PF01571; GCV_T; 1.
 CC Oxidoreductase; Flavoprotein; FAD; Mitochondrion; Transit peptide.
 CC

FT TRANSIT 1 43 MITOCHONDRION (POTENTIAL).
 FT CHAIN 44 857 DIMETHYLGLYCINE DEHYDROGENASE.
 FT BINDING 7 FAD (COVALENT).
 SQ SEQUENCE 857 AA; 96047 MW; A6FED946B8AB9878 CRC64;
 Query Match
 Best Local Similarity 71.1%; Score 32; DB 1; Length 857;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 2 FSAMGAE 8
 1 1 1 1 1
 DB 717 FRAMGSE 723
 RESULT 14
 CRG5_XENLA STANDARD: PRT; 132 AA.
 ID CRG5_XENLA
 AC 006255; -
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE GAMMA CRYSTALLIN V (GAMMA-5-CRY) (FRAGMENT).
 GN GCRY5.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 CC Xenopodinae; Xenopus.
 CC NCBI_TaxID=8355;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=93292986; PubMed=8514186;
 CC Smolich B.D., Tarrington S.K., Saha M.S., Stathakis D.G.,
 CC Gralinger R.M.;
 CC "Characterization of Xenopus laevis gamma-crystallin-encoding genes";
 CC RT Gene 128:189-195(1993)
 CC -1- FUNCTION: CRYSTALLINS ARE THE DOMINANT STRUCTURAL COMPONENTS
 CC OF THE VERTEBRATE EYE LENS.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- DOMAIN: HAS A TWO-DOMAIN BETA-STRUCTURE, FOLDED INTO FOUR
 CC VERY SIMILAR GREEK KEY MOTIFS.
 CC -1- SIMILARITY: BELONGS TO THE BETA/GAMMA-CRYSTALLIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M99583; AAA49694.1; -
 CC HSSP: P02526; IGCS.
 CC InterPro: IPR001064; -
 CC DR Pfam: PF00030; Crystallin; 1.
 CC PROSITE: PS00225; CRYSTALLIN, BETA/GAMMA; 2.
 CC Eye lens protein; Duplication; Multigene family.
 CC FT NON_TER 1 1 MOTIF 2.
 CC FT DOMAIN 42 45 CONNECTING PEPTIDE.
 CC FT DOMAIN 46 86 MOTIF 3.
 CC FT DOMAIN 87 132 MOTIF 4.
 CC SEQUENCE 132 AA; 16461 MW; CBA97C33E2F75377 CRC64;
 SQ
 Query Match
 Best Local Similarity 68.9%; Score 31; DB 1; Length 132;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 FSAMGAE 7
 1 1 1 1 1
 DB 111 RYSDMGA 117


```

RESULT 15
EV_EIAV9 STANDARD: PRT: 135 AA.
D REV_EIAV9
C P11305:
T 01-JUL-1989 (Rel. 11, Created)
T 01-JUL-1989 (Rel. 11, Last sequence update)
T 01-FEB-1996 (Rel. 33, Last annotation update)
E REV PROTEIN (3'-ORF PROTEIN) (FRAGMENT).
N REV.
S Equine infectious anemia virus (clone 1369) (EIAV).
C Viruses; Retroviral viruses; Retroviridae; Lentivirus.
C NCBI_TaxID:11670;
N [1]
P SEQUENCE FROM N.A.
X MEDLINE-87236196; PubMed-3035786;
A Kawakami T., Sherman L., Dahlberg J., Gazit A., Yaniv A.,
A Tronick S.R., Aaronson S.A.;
* Nucleotide sequence analysis of equine infectious anemia virus
  proviral DNA.
L Virology 158:300-312(1987).
C -1- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
C NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
C -1- SUBCELLULAR LOCATION: NUCLEAR PROTEIN THAT ACCUMULATES IN THE
C NUCLEOLI.
C -----
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@sib-sib.ch).
C -----
C EMBL: M16575; AAB59864.1; ALT_SEQ.
R PIR: D27842; ASLJEM.
W Transcription regulation; Trans-acting factor; Activator;
W Nuclear protein.
F NON_TER 1
T SEQUENCE 135 AA; 15988 MW; 8C56A9E96CD85DCD CRC64;
Q

```

Query Match 68.9%; Score 31; DB 1; Length 135;
 Best Local Similarity 100.0%; Pred.No. 43;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 FSAMG 6
 |||||
 b 100 FSAMG 104

Search completed: October 18, 2001, 16:52:30
 Job time: 115 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2001, 16:50:35 ; Search time 33.31 seconds
(without alignments)
31.775 Million cell updates/sec

Title: US-09-646-579-1
Perfect score: 45
Sequence: 1 RFSAWGAE 8

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

num DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	80.0	500	2	P96264
2	35	77.8	34	2	Q9F2X2
3	35	77.8	433	2	Q9L270
4	35	77.8	536	2	Q9H208
5	35	77.8	658	2	Q9KZ09
6	34	75.6	357	2	Q9RXX3
7	34	75.6	454	10	Q9FLE1
8	34	75.6	545	1	Q9V283
9	34	75.6	740	4	Q9U001
10	34	75.6	745	11	Q54697
11	33	73.3	272	2	Q9HXR8
12	33	73.3	386	2	Q9R097
13	33	73.3	430	2	Q9S3X4
14	33	73.3	432	2	Q9L629
15	33	73.3	575	10	Q9FTT6
16	33	73.3	635	14	Q69087
17	33	73.3	636	14	Q69104
18	33	73.3	637	2	Q44233
19	33	73.3	637	14	P89449

20	33	73.3	638	14	Q69527
21	33	73.3	707	8	Q98715
22	33	73.3	1130	2	Q9K2K8
23	33	73.3	1134	2	Q9L908
24	33	73.3	1150	2	Q53637
25	32	71.1	101	2	Q9RBR7
26	32	71.1	128	14	Q88533
27	32	71.1	163	4	Q9H7F5
28	32	71.1	171	14	Q69029
29	32	71.1	216	2	Q9RC08
30	32	71.1	272	5	Q9XU07
31	32	71.1	311	1	Q26780
32	32	71.1	383	10	Q82375
33	32	71.1	413	5	Q9VJPI
34	32	71.1	466	2	Q9K2C7
35	32	71.1	478	5	Q9NK87
36	32	71.1	481	2	Q87729
37	32	71.1	481	2	Q9KVU7
38	32	71.1	485	2	Q9RVV5
39	32	71.1	509	2	Q9RNP1
40	32	71.1	601	2	P71509
41	32	71.1	633	2	Q24759
42	32	71.1	636	2	Q9X760
43	32	71.1	649	5	Q9VBP7
44	32	71.1	796	2	Q9RL24
45	32	71.1	797	2	Q9L024

ALIGNMENTS

RESULT 1

P96264 PRELIMINARY: PRT; 500 AA.

AC P96264;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 13, Last annotation update)
DE HYPOTHETICAL. 52.0 KDA PROTEIN.
GN RV0418 OR MTCY22610.15.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S., Hovansky T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL: Z84724; CAB06558.1; -.
DR HSSP: P80561; IXYO.
DR Tuberculin; RV0418; -.
DR InterPro: IPR003137; -.
DR Pfam: PF02225; PA; 1.
KW Hypothetical protein.
SQ SEQUENCE 500 AA; 52042 MW; CD19BF47D8DB25EF CRC64;

Query Match 80.0%; Score 36; DB 2; Length 500;
Best Local Similarity 75.0%; Pred. No. 82;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 RFSAWGAE 8

11: 1111
b 309 RFAWGA 316

RESULT 2
PRELIMINARY; PRT; 324 AA.

09F2X2; 01-MAR-2001 (TREMblrel. 16, Created)

01-MAR-2001 (TREMblrel. 16, Last sequence update)

01-MAR-2001 (TREMblrel. 16, Last annotation update)

PURATIVE AMINOPEPTIDASE.

SCD20.07C.

Streptomyces coelicolor.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.

NCBI_TaxID=1902;

SEQUENCE FROM N.A.

STRAIN-A3(2);

Saunders D.C., Harris D.,

Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.,

Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

MEDLINE=97000351; PubMed=8843436;

Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,

Kinashi H., Hopwood D.A.,

"A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

Mol. Microbiol. 21:77-96(1996).

EMBL: AL392148; CAC08290.1; -

AMINOPEPTIDASE.

RESULT 4
PRELIMINARY; PRT; 536 AA.

09H208; 01-MAR-2001 (TREMblrel. 16, Created)

01-MAR-2001 (TREMblrel. 16, Last sequence update)

01-MAR-2001 (TREMblrel. 16, Last annotation update)

PROBABLE AMINOPEPTIDASE.

PA2939.

Pseudomonas aeruginosa.

Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

Pseudomonas.

NCBI_TaxID=287;

SEQUENCE FROM N.A.

STRAIN-PA01;

MEDLINE=20437337; PubMed=10984043;

Stover C.K., Pham X.-Q.T., Ervin A.L., Mizoguchi S.D., Warren P.,

Hickey R.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Ryan Y.,

Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,

Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

"Complete genome sequence of Pseudomonas aeruginosa PA01, an

opportunistic pathogen.";

Nature 406:959-964(2000).

EMBL: AE004720; AAG06327.1; -

InterPro: IPR003137; -

DR EMBL: AE004720; AAG06327.1; -

DR InterPro: IPR003137; -

DR Pfam: PF02225; PA: 1.

AMINOPEPTIDASE.

SEQUENCE 536 AA; 57511 MW; EE57BEFFC1C803A CRC64;

RESULT 5
PRELIMINARY; PRT; 658 AA.

09KZ08; 01-OCT-2000 (TREMblrel. 15, Created)

01-OCT-2000 (TREMblrel. 15, Last sequence update)

01-OCT-2000 (TREMblrel. 15, Last annotation update)

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Seeger K.J., Harris D.,

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Seeger K.J., Harris D.,

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Seeger K.J., Harris D.,

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Seeger K.J., Harris D.,

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Seeger K.J., Harris D.,

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN-A3(2);

Seeger K.J., Harris D.,

Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DE ACETOACETYL-COA SYNTHETASE.
 GN ACSA.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 RX NCBI_TaxID:1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Harris D.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-97000351; PubMed-8843436;
 RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
 P Kinashi H., Hopwood D.A.;
 P "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL353861; CAB88889.1; -;
 DR InterPro: IPR000873; -;
 DR PROSITE: PS00455; AMP_BINDING; 1.
 SQ SEQUENCE 658 AA; 72137 MW; 34AF9A24FF22F11D CRC64;

Query Match 77.8%; Score 35; DB 2; Length 658;
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RFSAMGA 8
 111111
 DB 24 RFSAMGA 31

RESULT 6
 ID Q9RX3 PRELIMINARY; PRT; 357 AA.
 AC Q9RX3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 RN DR2359.
 GN Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 RX NCBI_TaxID:1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RI.
 RL MEDLINE-20036896; PubMed-10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 Matarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 Fraser C.M.;
 P "Genome sequence of the radioresistant bacterium Deinococcus
 radiodurans R1."
 RL Science 286:1571-1577(1999).
 DR EMBL: AE002066; AAF11904.1; -;
 DR TIGR: DR2359; -;
 SQ SEQUENCE 357 AA; 40102 MW; 17A7D864C6184081 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 357;
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RFSAMGA 8
 111111
 DB 123 RFSAMGA 130

RESULT 7
 ID Q9FLE1 PRELIMINARY; PRT; 454 AA.
 AC Q9FLE1;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GENOMIC DNA. CHROMOSOME 5, P1 CLONE:MH19.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 RX NCBI_TaxID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RL MEDLINE-98290546; PubMed-9628582;
 RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
 RA Tabata S.;
 P "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 Physically assigned p1 and TAC clones."
 RL DNA Res. 5:41-54(1998).
 DR EMBL: AB010077; BAB10218.1; -;
 SQ SEQUENCE 454 AA; 48839 MW; EF503DA8BD24F730 CRC64;

Query Match 75.6%; Score 34; DB 10; Length 454;
 Best Local Similarity 85.7%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RFSAMGA 7
 111111
 DB 295 RFSAMGA 301

RESULT 8
 ID Q9V283 PRELIMINARY; PRT; 545 AA.
 AC Q9V283;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE SUGAR ABC TRANSPORTER BINDING PROTEIN RELATED.
 GN PAB2439.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococcales; Pyrococcus.
 RX NCBI_TaxID:29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORSAY;
 RL Heilig R.;
 P "Pyrococcus abyssi genome sequence: Insights into archaeal chromosome
 structure and evolution."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ248283; CAB49115.1; -;
 DR InterPro: IPR000567; -;
 DR Pfam: PF01547; SBP_bacterial_1; 1.
 SQ SEQUENCE 545 AA; 61787 MW; 6ED118EE38FB5C34 CRC64;

Query Match 75.6%; Score 34; DB 1; Length 545;
 Best Local Similarity 71.4%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RFSAMGA 7

b 63 KESAMGS 69

RESULT 9

D 09UC01 PRELIMINARY; PRT: 740 AA.
 C 09UC01: 01-MAY-2000 (TREMBLrel. 13, Created)
 T 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 T 01-OCT-2000 (TREMBLrel. 13, Last annotation update)
 E NAALADASE L. PROTEIN.
 N NAALADASE L.
 S Homo sapiens (Human).
 C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 X NCBI_TaxID=9606;
 N [1]
 SEQUENCE FROM N.A.
 TISSUE=SMALL INTESTINE;
 MEDLINE=99185063; PubMed=10085079;
 A Pangalos M.N., Neefs J.M., Somers M., Fraiponts E., Ashton D.,
 Gordon R.G.;
 T "Isolation and expression of novel human glutamate carboxypeptidases
 with N-acetylated alpha-linked acidic dipeptidase and dipeptidyl
 peptidase IV activity.";
 L J. Biol. Chem. 274:8470-8483(1999).
 R EMBL: A012371; CAB39968.1; -;
 R InterPro: IPR000719; -;
 R InterPro: IPR003137; -;
 R Pfam: PF02225; PA: 1.
 R PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN 1.
 O SEQUENCE 740 AA; 80620 MW; E2E5AAC9E5056A14 CRC64;

Query Match 75.6%; Score 34; DB 4; Length 740;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 2 F5AMGAE 8
 I:::||||
 b 410 F5AMGAE 416

RESULT 10
 054697 PRELIMINARY; PRT: 745 AA.
 054697:

T 01-JUN-1998 (TREMBLrel. 06, Created)
 T 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 T 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 E 1100.
 S Rattus norvegicus (Rat).
 C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 X NCBI_TaxID=10116;
 N [1]
 SEQUENCE FROM N.A.
 C STRAIN=SPRAGUE-DAWLEY; TISSUE=ILEUM;
 A Schneider B.L., Thevananthar S., Moyer M.S., Walters H.C., Rhinado P.,
 Devaiajan P., Sun A.O., Dawson P.A., Ananthanarayanan M.;
 L J. Biol. Chem. 0:0-0(1997).
 R EMBL: AF009921; AAB87644.1; -;
 R MEROPS: M28.011; -;
 R InterPro: IPR003137; -;
 R Pfam: PF02225; PA: 1.
 O SEQUENCE 745 AA; 80640 MW; A59C2EFD23BE36B5 CRC64;

Query Match 75.6%; Score 34; DB 11; Length 745;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 F5AMGAE 8
 I:::||||
 Db 415 F5AMGAE 421

RESULT 11

ID 09HX80 PRELIMINARY; PRT: 272 AA.
 AC 09HX80:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE PROBABLE PERMEASE OF ABC TAURINE TRANSPORTER.
 GN PA3936.
 GN Pseudomonas aeruginosa.
 OS Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 X NCBI_TaxID=287;
 N [1]
 SEQUENCE FROM N.A.
 RC STRAIN=PA01;
 RA MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Madman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RT Nature 406:959-964(2000).
 RL EMBL: AE004811; AAG07323.1; -;
 DR InterPro: IPR000515; -;
 DR Pfam: PF00528; BPD_transp; 1.
 SO SEQUENCE 272 AA; 29447 MW; D64898B300B2D06F CRC64;

Query Match 73.3%; Score 33; DB 2; Length 272;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 F5AMGAE 6
 I:::||||
 Db 267 F5AMGAE 272

RESULT 12

ID 09R097 PRELIMINARY; PRT: 386 AA.
 AC 09R097:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE CYSTATHIONINE GAMMA-SYNTHASE.
 GN METB.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
 OC Corynebacterium.
 X NCBI_TaxID=1718;
 N [1]
 SEQUENCE FROM N.A.
 RC STRAIN=AS019;
 RX MEDLINE=99347339; PubMed=10420990;
 RA Hwang B.J., Kim Y., Kim H.-B., Hwang H.J., Kim J.H., Lee H.S.;
 RT "Analysis of corynebacterium glutamicum methionine biosynthetic
 pathway: isolation and analysis of metB encoding cystathionine gamma-
 synthase.";
 RT Mol. Cells 9:300-308(1999).
 RL EMBL: AF126953; AAD54070.1; -;
 DR InterPro: IPR000277; -;
 DR Pfam: PF01053; Cys_Met_Meta_PP; 1.
 SO SEQUENCE 386 AA; 41658 MW; 72EBB2B34275296A CRC64;

Query Match 73.3%; Score 33; DB 2; Length 386;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FSAWGAE 8
11111111
DB 118 FTAWGVE 124

RESULT 13

Q9S3X4 PRELIMINARY; PRT; 430 AA.
AC Q9S3X4; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, last annotation update)
DE SPT.
GN SPT.
OS Streptomyces spectabilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=68270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NRRL 2494;
RA Altenbuchner J., Jennen D., Voff J.N.;
RT "The spectinomycin resistance gene of the spectinomycin producer
Streptomyces spectabilis."
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF10704; AAD50454.1;
SQ SEQUENCE 430 AA; 44966 MW; 91DF225A21906AAA CRC64;

Query Match 73.3%; Score 33; DB 2; Length 430;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RFSAMGA 7
11111111
DB 367 RFAWGGA 373

RESULT 14

Q9L629 PRELIMINARY; PRT; 432 AA.
AC Q9L629; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE EXPORT PROTEIN.
GN SPT.
OS Streptomyces spectabilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=68270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC27741;
RA Song J.K., Kim S.B., Cha J.H., Hahn J.J., Kim J.W., Hyun C.G.,
Sub J.W.;
RT "Cloning and Identification of a Gene Cluster in Streptomyces
spectabilis Spectinomycin Producer."
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF244574; AAF63340.1;
SQ SEQUENCE 432 AA; 45316 MW; 99B4FC1B822EBBC0 CRC64;

Query Match 73.3%; Score 33; DB 2; Length 432;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RFSAMGA 7
11111111
DB 369 RFAWGGA 375

RESULT 15

Q9FTI6 PRELIMINARY; PRT; 575 AA.
AC Q9FTI6; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE P0410E01.14 PROTEIN.
GN P0410E01.14.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
OC Oryza.
GN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0410E01."
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP002866; BABI7093.1;
SQ SEQUENCE 575 AA; 60463 MW; D100FB76FA4A4C134 CRC64;

Query Match 73.3%; Score 33; DB 10; Length 575;
Best Local Similarity 71.4%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RFSAMGA 7
11111111
DB 303 RFAWGGA 309

Search completed: October 18, 2001, 16:53:10
Job time: 155 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2001, 16:51:16 ; Search time 32.41 Seconds
(Without alignments)
16.835 Million cell updates/sec

Title: US-09-646-579-2
Perfect score: 49
Sequence: 1 ISRFAMGEV 9

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 segs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
mun DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.0601.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	9	AAW56752	A. calcoaceticus a
2	49	100.0	9	AAW42322	Actinobacter-deri
3	35	71.4	9	AAW1853	MBP peptide 110, p
4	35	71.4	10	AAW1951	MBP peptide 110, p
5	35	71.4	11	AAW1960	MBP peptide 108, p
6	35	71.4	12	AAW27077	Myelin basic prote
7	35	71.4	13	AAW4964	MHC binding peptid
8	35	71.4	19	AAW5345	MBP-3.1 (111-129).
9	-35	71.4	19	AAW43951	Human myelin basic
10	35	71.4	20	AAW5394	Residues 101-120 o
11	35	71.4	20	AAW5344	MBP-3 (111-130).

12	35	71.4	20	18	AAW43950	Human myelin basic
13	35	71.4	20	18	AAW44066	Human myelin basic
14	35	71.4	20	20	AAW73618	Human myelin basic
15	35	71.4	20	21	AAW12619	Human myelin basic
16	35	71.4	20	21	AAW85551	Human myelin basic
17	35	71.4	25	17	AAW95367	Human MBP peptide
18	35	71.4	25	17	AAW95367	Residues 111-135 o
19	35	71.4	25	18	AAW43953	MBP-5 (101-125).
20	35	71.4	127	15	AAW48596	Human myelin basic
21	35	71.4	167	15	AAW48595	Rat myelin basic p
22	35	71.4	168	11	AAW04717	Myelin basic prote
23	35	71.4	168	15	AAW48594	Empirically determ
24	35	71.4	169	15	AAW48593	Rabbit myelin basi
25	35	71.4	170	11	AAW02226	Cattle myelin basi
26	35	71.4	170	14	AAW30735	Myelin basic prote
27	35	71.4	170	14	AAW30736	Bovine MBP. Synth
28	35	71.4	170	14	AAW35440	Human MBP. Synth
29	35	71.4	170	15	AAW45947	Human basic myelin
30	35	71.4	170	15	AAW48592	Human basic myelin
31	35	71.4	170	17	AAW48596	Human myelin basic
32	35	71.4	170	17	AAW57236	Myelin oligodendro
33	35	71.4	170	19	AAW86422	Myelin basic prote
34	35	71.4	170	17	AAW58368	Myelin basic prote
35	35	71.4	171	17	AAW97627	Mammalian generic
36	35	71.4	171	17	AAW9580	Human myelin basic
37	35	71.4	171	20	AAW12784	Human myelin basic
38	35	71.4	171	21	AAW19513	Human myelin basic
39	35	71.4	171	21	AAW69394	Human myelin basic
40	35	71.4	171	21	AAW44234	Human myelin basic
41	35	71.4	197	17	AAW00399	Amino acid sequenc
42	35	71.4	197	21	AAW9582	Human myelin basic
43	35	71.4	203	17	AAW06107	Human myelin bindi
44	35	71.4	203	17	AAW06108	Foetal myelin basi
45	35	71.4	373	17	AAW06103	MP4 chimera (MBP21

ALIGNMENTS

RESULT	1
AAW56752	standard; peptide; 9 AA.
ID	AAW56752;
AC	31-JUL-1998 (first entry)
DT	
DE	
DE	A. calcoaceticus antigenic peptide exhibiting molecular mimicry.
KW	BSE, bovine; myelin; molecular mimicry; antigenic; assay; diagnosis;
KW	spongiform encephalopathy; Creutzfeldt Jacob disease; multiple sclerosis;
KW	Actinobacter calcoaceticus; Agrobacterium; Rhinococcus.
OS	Actinobacter calcoaceticus.
OS	
PN	W09813694-A1.
XX	
XX	02-APR-1998.
XX	
XX	29-SEP-1997; 97WO-GB02667.
XX	
XX	27-SEP-1996; 96GB-0020195.
XX	
XX	(UNIL) KING'S COLLEGE.
XX	
XX	Ehringer A;
XX	
XX	WPI: 1998-230844/20.
XX	
XX	Diagnosis of demyelinating conditions, particularly BSE - by
XX	assaying for antibodies which bind to an antigenic peptide which
XX	exhibits molecular mimicry of a mammalian myelin peptide

15 Claim 7; Page 7; 12pp; English.

16 This is a Acinetobacter calcoaceticus antigenic peptide that exhibits

17 molecular mimicry to a bovine myelin peptide. This is used in the methods

18 and diagnostic tests for spongiform encephalopathy (SE) and other

19 demyelinating conditions in mammals which comprises assaying antibodies

20 present in the mammal which bind to an antigenic peptide which exhibits

21 molecular mimicry of a mammalian myelin peptide. A method for testing for

22 bovine SE (BSE) in cattle comprises assaying sera collected from the

23 cattle for antibodies to a species of Acinetobacter. Agrobacterium or

24 Rhinococcus, or a peptide having a sequence present in the species which

25 mimics a peptide of bovine myelin and identifying animals having a level

26 of antibodies at least 2 standard deviations above that of healthy

27 control animals. The diagnostic test kit for BSE in cattle comprises as

28 test antigen a species of Acinetobacter, Rhinococcus or Agrobacterium or

29 a peptide having a sequence present in the species which mimics a peptide

30 of bovine myelin. The methods can be used for diagnosis of demyelinating

31 conditions such as BSE, Creutzfeldt Jakob disease and multiple sclerosis.

32 Sequence 9 AA;

33

34 Query Match 100.0%; Score 49; DB 19; Length 9;

35 Best Local Similarity 100.0%; Pred. No. 3.4e+05;

36 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

37

38 1 ISRFAMGEV 9

39 | | | | | | | | |

40 1 Isrfawgev 9

41

42 RESULT 2

43 AY42322

44 AAY42322 standard; peptide; 9 AA.

45

46 AAY42322;

47

48 06-DEC-1999 (first entry)

49

50 Acinetobacter-derived peptide epitope #2.

51

52 Acinetobacter; myelin; spongiform encephalopathy; CJD; immune response;

53 Creutzfeldt-Jakob disease; multiple sclerosis; antibody; epitope;

54 detection.

55

56 Synthetic.

57 Acinetobacter sp.

58

59 WO947932-A2.

60

61 23-SEP-1999.

62

63 19-MAR-1999; 99MO-GB00876.

64

65 19-MAR-1998; 98GB-0005913.

66

67 (UNLO) UNIV LONDON KING'S COLLEGE.

68

69 Ehringer A;

70

71 WPI; 1999-571874/48.

72

73 Detecting de-myelinating disease or spongiform encephalopathy

74

75 Claim 12; Page 7; 11pp; English.

76

77 This sequence represents a synthetic peptide epitope (#2) derived from

78 species of Acinetobacter, which is sufficiently conformationally similar

79 to natural Acinetobacter peptide epitopes to bind the corresponding

80 antibodies. Acinetobacter species may be implicated in spongiform

81 encephalopathies such as Creutzfeldt-Jakob disease or in de-myelinating

82 diseases such as multiple sclerosis. Involvement of Acinetobacter forms

83 the basis for a model for the development of de-myelinating diseases (and

CC an alternative model for the development of spongiform encephalopathies).

CC This model involves the phenomenon of molecular mimicry in which mammals

CC exposed to certain bacteria containing peptide sequences which mimic

CC myelin peptides, experience an autoimmune response. A novel method for

CC detecting a de-myelinating disease or spongiform encephalopathy in

CC mammals has been developed which comprises testing a biological sample

CC obtained from the mammal for immunoglobulin A (IgA) antibodies indicative

CC of infection by Acinetobacter species. This sequence can be used as a

CC test antigen in a kit to detect such antibodies. The method is useful for

CC diagnosing de-myelinating diseases and spongiform encephalopathies in

CC animals and humans, especially bovine spongiform encephalopathy, multiple

CC sclerosis or Creutzfeldt-Jacob disease in humans. The method is useful

CC for the early detection of these infections, resulting in early

CC treatments.

CC

CC Sequence 9 AA;

CC

CC Query Match 100.0%; Score 49; DB 20; Length 9;

CC Best Local Similarity 100.0%; Pred. No. 3.4e+05;

CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC

CC 1 ISRFAMGEV 9

CC | | | | | | | | |

CC 1 Isrfawgev 9

CC

CC RESULT 3

CC AAR61853

CC AAR61853 standard; peptide; 9 AA.

CC

CC AAR61853;

CC

CC 18-MAY-1995 (first entry)

CC

CC MBP peptide 110, potential binder of HLA-A2.1.

CC

CC antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV;

CC HIV1; plasma specific antigen; hepatitis B virus; Epstein Barr;

CC human immunodeficiency virus; human papilloma virus; p53; c-ERB2;

CC MAGE-1; melanoma antigen-1; core antigen; surface antigen;

CC pharmaceutical composition; in vivo; ex vivo; therapeutic;

CC diagnostic; MHC class I molecule; major histocompatibility complex;

CC HLA-A2.1; 9mer; 10mer; anchor; human leukocyte antigen; PLP; 8mer;

CC algorithm prediction; MBP.

CC

CC Homo sapiens.

CC

CC WO9420127-A.

CC

CC 15-SEP-1994.

CC

CC 04-MAR-1994; 94MO-US02353.

CC

CC 05-MAR-1993; 93US-0027146.

CC 04-JUN-1993; 93US-0073205.

CC 29-NOV-1993; 93US-0159184.

CC

CC (CYTE-) CYTEL CORP.

CC

CC Grey HM, Kast WM, Sette A, Sidney J;

CC

CC WPI; 1994-302678/37.

CC

CC Immunogenic peptide(s) having an HLA-A2.1 binding motif - used

CC for treatment or prophylaxis of cancer, virus infection or

CC autoimmune diseases.

CC

CC Disclosure; Page 122; 138pp; English.

CC

CC AAR61714-837 are potential peptide binders of HLA-A2.1 motif. These

CC peptides are thus potentially immunogenic. They were predicted by

CC using an algorithm, which assigns a score for each amino acid, at

CC each position along a peptide. A peptide is scored in the 'Grouped
CC Ratio' algorithm as a product of the scores of each of its residues.
CC This peptide has an algorithm score (EO2) of -21.42. This value can
CC then be used to predict a population of peptides with the highest
CC occurrence of good binders. The peptides of the invention can induce
CC cytotoxic T lymphocytes which can react with target cells. They can
CC be used for the treatment or prophylaxis of cancer, eg. prostate
CC cancer or lymphoma, etc.

CC
XX
SQ Sequence 9 AA;

Query Match 71.4%; Score 35; DB 15; Length 9;
Best Local Similarity 71.4%; Pred. No. 3.4e+05;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
Db :|||:|
2 lstrswg 8

RESULT 4
A*61951

AA61951 standard; peptide; 10 AA.

AC AAR61951;

DT 31-MAY-1995 (first entry)

DE MBP peptide 110, potential binder of HLA-A2.1.

antigen; epitope; immunogenic target protein; PSA; HBVC; HBVs; EBV;
HIV1; plasma specific antigen; hepatitis B virus; Epstein Barr;
human immunodeficiency virus; human papilloma virus; p53; c-ERB2;
MAGE-1; melanoma antigen-1; core antigen; surface antigen;
pharmaceutical composition; in vivo; ex vivo; therapeutic;
diagnostic; MHC class I molecule; major histocompatibility complex;
HLA-A2.1; 9mer; 10mer; anchor; human leukocyte antigen; PLP; 8mer;
algorithm prediction; MBP.

OS Homo sapiens.

PN MO9420127-A.

PD 15-SEP-1994.

PF 04-MAR-1994; 94MO-US02353.

PR 05-MAR-1993; 930S-0027146.

PR 04-JUN-1993; 930S-0073205.

PR 29-NOV-1993; 930S-0159184.

PT (CYTE-) CYTEL CORP.

PT Grey HM, Kast WM, Sette A, Sidney J;

PT WPI; 1994-302678/37.

Immunogenic peptide(s) having an HLA-A2.1 binding motif - used
for treatment or prophylaxis of cancer, virus infection or
autoimmune diseases.

Disclosure; Page 128; 138pp; English.

AA61714-837 are potential peptide binders of HLA-A2.1 motif. These
peptides are thus potentially immunogenic. They were predicted by
using an algorithm, which assigns a score for each amino acid, at
each position along a peptide. A peptide is scored in the 'Grouped
Ratio' algorithm as a product of the scores of each of its residues.
This score has an algorithm score (EO2) of -28.38. This value can
then be used to predict a population of peptides with the highest
occurrence of good binders. The peptides of the invention can induce
cytotoxic T lymphocytes which can react with target cells. They can

CC be used for the treatment or prophylaxis of cancer, eg. prostate
CC cancer or lymphoma, etc.

CC
XX
SQ Sequence 10 AA;

Query Match 71.4%; Score 35; DB 15; Length 10;
Best Local Similarity 71.4%; Pred. No. 2.5;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
Db :|||:|
2 lstrswg 8

RESULT 5
AAR61960

ID AAR61960 standard; peptide; 11 AA.

AC AAR61960;

DT 31-MAY-1995 (first entry)

DE MBP peptide 108, potential binder of HLA-A2.1.

antigen; epitope; immunogenic target protein; PSA; HBVC; HBVs; EBV;
HIV1; plasma specific antigen; hepatitis B virus; Epstein Barr;
human immunodeficiency virus; human papilloma virus; p53; c-ERB2;
MAGE-1; melanoma antigen-1; core antigen; surface antigen;
pharmaceutical composition; in vivo; ex vivo; therapeutic;
diagnostic; MHC class I molecule; major histocompatibility complex;
HLA-A2.1; 9mer; 10mer; anchor; human leukocyte antigen; PLP; 8mer;
algorithm prediction; MBP.

OS Homo sapiens.

PN MO9420127-A.

PD 15-SEP-1994.

PF 04-MAR-1994; 94MO-US02353.

PR 05-MAR-1993; 930S-0027146.

PR 04-JUN-1993; 930S-0073205.

PR 29-NOV-1993; 930S-0159184.

PT (CYTE-) CYTEL CORP.

PT Grey HM, Kast WM, Sette A, Sidney J;

PT WPI; 1994-302678/37.

Immunogenic peptide(s) having an HLA-A2.1 binding motif - used
for treatment or prophylaxis of cancer, virus infection or
autoimmune diseases.

Disclosure; Page 129; 138pp; English.

AA61714-837 are potential peptide binders of HLA-A2.1 motif. These
peptides are thus potentially immunogenic. They were predicted by
using an algorithm, which assigns a score for each amino acid, at
each position along a peptide. A peptide is scored in the 'Grouped
Ratio' algorithm as a product of the scores of each of its residues.
This value can then be used to predict a population of peptides with
the highest occurrence of good binders. The peptides of the invention
can induce cytotoxic T lymphocytes which can react with target cells.
They can be used for the treatment or prophylaxis of cancer, eg.
prostate cancer or lymphoma, etc.

SO Sequence 11 AA;

Query Match 71.4%; Score 35; DB 15; Length 11;

Best Local Similarity 71.4%; Pred. No. 2.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 ISRFAMG 7
:|||||
b 4 lsrfswg 10

RESULT 6
AY27077
ID AAY27077 standard; peptide; 12 AA.

AY27077;

18-OCT-1999 (first entry)

Myelin basic protein (MBP) fragment (residues 110-121).

Inflammatory arthritis; antiarthritic; myelin basic protein; MBP; arthritis; vaccine.

Sus sp.

US5942491-A.

24-AUG-1999.

22-OCT-1996; 96US-0735253.

22-OCT-1996; 96US-0735253.

(UNMS) UNIV MICHIGAN STATE.

Root-Bernstein RS;

WPI; 1999-493521/41.

Treating inflammatory arthritis using peptides with homology to known proteins

Disclosure; Column 2: 23pp; English.

The invention relates to a method for treating inflammatory arthritis by administering antiarthritic peptides comprising an amino acid residue sequence that is identical to or homologous to a contiguous stretch of at least 10 amino acids of sequences shown in AAY27077 and AAY27078 (residues 110-121 or 153-162 of myelin basic protein (MBP)). The new method is useful for treating inflammatory arthritis in a mammal, especially a dog, horse or human, and for preventing the onset of arthritis. The peptides are useful as antiarthritic vaccines. The method specifically uses peptide sequences shown in AAY27085-92 for treating inflammatory arthritis. The present sequence represents a MBP peptide fragment.

Sequence 12 AA;

Query Match 71.4%; Score 35; DB 20; Length 12;
Best Local Similarity 71.4%; Pred. No. 3;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 ISRFAMG 7
:|||||
b 2 lsrfswg 8

RESULT 7
AM94964
D AAM94964 standard; peptide; 13 AA.

AAW94964;

12-MAY-1999 (first entry)

XX MHC binding peptide from myelin basic protein (MBP).
DE Major histocompatibility class II; MHC; binding; vaccine; MBP.
XX autoimmune disease; myelin basic protein; HLA-DRB1*0401.
KW
XX
XX
XX
OS Synthetic.

MO9859244-A1.

30-DEC-1998.

19-JUN-1998; 98WO-GB01801.

20-JUN-1997; 97GB-0012892.

(ECLA-) ECLAGEN LTD.

Brooks T, Carr F, Fothergill J, Kemp G;

WPI; 1999-105639/09.

Identifying MHC binding peptides - by ascertaining the characteristics of a MHC molecule binding groove and determining conformation scores of pocket bound peptides

Example 4; Page 50; 64pp; English.

The invention relates to a method for the prediction of the binding affinity of a peptide to a major histocompatibility (MHC) class II molecules. The method comprises (a) ascertaining the characteristics of a MHC molecule binding groove; (b) presenting a selected peptide to the MHC molecule and ascertaining a first conformation score for each pocket bound peptide side-chain; (c) amending the conformation of each pocket bound peptide side-chain and ascertaining a second conformation score; (d) repeating (c) with alternative conformations of each peptide pocket bound side-chain; (e) choosing the highest conformation score for each pocket bound peptide side-chain in each binding groove pockets; and (f) combining the highest conformation score for each pocket and ascertaining a binding score for the complete peptide. The invention also provides a computer conditioned to receive information characterizing a peptide bound to the MHC molecule and to utilise the information to predict the binding affinity of the peptide. The method can be used to identify peptides in a protein which can bind to a given MHC molecule. Such peptides can be used in the production of molecules for use in e.g. vaccines, autoimmune diseases and diagnostics. Sequences AAW94957 to AAW94961 represent peptides from myelin basic protein (MBP) that can bind with MHC molecule HLA-DRB1*0401.

Sequence 13 AA;

Query Match 71.4%; Score 35; DB 20; Length 13;
Best Local Similarity 71.4%; Pred. No. 3.3;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
:|||||
Db 5 lsrfswg 11

RESULT 8
AAR95345
ID AAR95345 standard; peptide; 19 AA.

AAR95345;

16-DEC-1996 (first entry)

MBP-3.1 (111-129).

Myelin basic protein; MBP; experimental allergic encephalomyelitis; EAE; CD4+; T-cell; autoimmune disease; demyelination; central nervous system;

CNS; animal model; human; multiple sclerosis; MS; mammal; benign MS; MOG relapsing-remitting MS; chronic progressive MS; myelin autoantigen; PLP; diabetes; Graves disease; myasthenia gravis; Good Pasture's syndrome; psoriasis; thyroiditis; rheumatoid arthritis; proteolipid protein; MAG; myelin oligodendrocyte protein; myelin associated glycoprotein; therapy.

Synthetic.

OS
XX
PN W09612737-A2.
XX
PD 02-MAY-1996.
XX
PF 25-OCT-1995; 95MO-US13682.
XX
PR 15-MAR-1995; 95US-0404228.
PR 25-OCT-1994; 94US-0328224.
PR 25-OCT-1995; 95ZA-0009033.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
PI Devaux B., Franzen H., Gefter M., Hsu D., Pallard X;
PI Rothbard J., Samson M., Shi J., Smillek D;
X* WPI: 1996-230552/23.

Myelin basic derived peptide(s) and analogs - used in the treatment of Multiple Sclerosis, psoriasis, Graves Disease, etc.

Claim 1; Fig 2; 91pp; English.

AAR95334-R95374 represent peptides derived from myelin basic protein (MBP). Immunisation with MBP can be used to induce experimental allergic encephalomyelitis (EAE) in susceptible strains of mice. EAE is a CD4+ T-cell mediated autoimmune disease which results in demyelination of the central nervous system, resulting in paralysis and other neurological abnormalities. EAE is a commonly used animal model for human multiple sclerosis (MS). These sequences can be used in compositions for treating MS in a mammal. The composition acts to down regulate the autoimmune response, and may be administered in an amount sufficient to prevent the onset of symptoms of MS. The compositions may also be used to treat advanced stage MS, especially relapsing-remitting MS, chronic progressive MS or benign MS. These peptides may also be used in the treatment of other diseases involving myelin autoantigens, including diabetes, Graves disease, myasthenia gravis, Good Pasture's syndrome, psoriasis, thyroiditis, and rheumatoid arthritis. Peptides derived from other myelin autoantigens, such as myelin oligodendrocyte protein (MOG), proteolipid protein (PLP), and myelin associated glycoprotein (MAG) can be used as alternatives to these MBP peptides in these compositions.

Sequence 19 AA:

.erry Match 71.4%; Score 35; DB 17; Length 19;
Best Local Similarity 71.4%; Pred. No. 4.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0.

OY 1 ISRRFNG 7
DB 1 lsrtswg 7

RESULT 9
ID AAM43951 standard; peptide: 19 AA.
AAM43951
AC
XX AAM43951;
XX
XT 22-APR-1998 (first entry)
DE Human myelin basic protein peptide MBP-3.1 (111-129).
KW Human; myelin oligodendrocyte glycoprotein; MOG; multiple sclerosis;
W autoimmune MBP; myelin basic protein; demyelinating.

```

XX OS Homo sapiens.
XX PN W09735879-A1.
XX PD 02-OCT-1997.
XX PF 01-MAY-1996; 96MO-US06072.
XX PR 28-MAR-1996; 96DS-0623406.
XX PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX PI Devaux B, Garman RD, Rothbard J, Smilek D, Wallner B;
XX WPI: 1997-489564/45.
XX DR
XX PT Human myelin oligodendrocyte glyco-protein peptide - useful to treat
XX PT multiple sclerosis in mammal by down-regulating auto-immune response
XX PS Claim 101; Page 35; 108pp; English.
XX SS
XX SQ The present sequence represents a novel peptide of human myelin
CC basic protein (MBP). The peptide of human MOG (myelin oligodendrocyte
CC glycoprotein) and MBP can be used to treat multiple sclerosis in a
CC mammal by down-regulating an autoimmune response in the mammal. They
CC can also be used to diagnose and treat other demyelinating autoimmune
CC diseases in humans, or to prepare antibodies for the detection or
CC diagnosis of autoimmune diseases.
XX SQ Sequence 19 AA:

Query Match 71.4%; Score 35; DB 18; Length 19;
Best Local Similarity 71.4%; Pred. No. 4.8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0.
QY 1 ISRFWNG 7
   :|||:|
Db 1 ISrlswg 7

RESULT 10
AAR95394
ID AAR95394 standard; peptide; 20 AA.
AC AAR95394;
AD
DE Residues 101-120 of myelin basic protein.
DX
DY 16-DEC-1996 (first entry)
EY
FZ
GZ
HZ
IY
JY
KZ
LZ
MY
NZ
OZ
PZ
QZ
RZ
SZ
TZ
UZ
VZ
WZ
XZ
YZ
ZZ
AA (IMMU-) IMMUNOLOGIC PHARM CORP.

```

Devauux B, Franzen H, Geffer M, Hsu D, Pallard X;
 Rothbard J, Samson M, Shi J, Smilek D;
 WPI; 1996-230552/23.

Myelin basic derived peptide(s) and analogs - used in the treatment
 of Multiple Sclerosis, psoriasis, Graves Disease, etc.

Example 1; Fig 3; 91pp; English.

AA95334-R95374, and AA95387-R95397 represent peptides derived from
 myelin basic protein (MBP). Immunisation with MBP can be used to induce
 experimental allergic encephalomyelitis (EAE) in susceptible strains of
 mice. EAE is a CD4+ T-cell mediated autoimmune disease which results in
 demyelination of the central nervous system, resulting in paralysis and
 other neurological abnormalities. EAE is a commonly used animal model
 for human multiple sclerosis (MS). These sequences can be used in
 compositions for treating MS in a mammal. The composition acts to down
 regulate the autoimmune response, and may be administered in an amount
 sufficient to prevent the onset of symptoms of MS. The compositions may
 also be used to treat advanced stage MS, especially relapsing-remitting
 MS, chronic progressive MS or benign MS. These peptides may also be
 used in the treatment of other diseases involving myelin autoantigens,
 including diabetes, Graves disease, myasthenia gravis, Good Pasture's
 syndrome, psoriasis, thyroiditis, and rheumatoid arthritis. Peptides
 derived from other myelin autoantigens, such as myelin oligodendrocyte
 protein (MOG), proteolipid protein (PLP), and myelin associated
 glycoprotein (MAG) can be used as alternatives to these MBP peptides in
 these compositions.

Sequence 20 AA:

Query Match 71.4%; Score 35; DB 17; Length 20;
 Best Local Similarity 71.4%; Pred. No. 5;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 ISRFAMG 7
 :|||:|
 11 Isrlswg 17

RESULT 11

AA95344 standard; peptide; 20 AA.

AA95344;

16-DEC-1996 (first entry)

MBP-3 (111-130).

Myelin basic protein; MBP; experimental allergic encephalomyelitis; EAE;
 CD4+ T-cell; autoimmune disease; demyelination; central nervous system;
 CNS; animal model; human; multiple sclerosis; MS; mammal; benign MS; MOG;
 relapsing-remitting MS; chronic progressive MS; myelin autoantigen; PLP;
 diabetes; Graves disease; myasthenia gravis; Good Pasture's syndrome;
 psoriasis; thyroiditis; rheumatoid arthritis; proteolipid protein; MAG;
 myelin oligodendrocyte protein; myelin associated glycoprotein; therapy.

Synthetic.

WO9612737-A2.

02-MAY-1996.

25-OCT-1995; 95MO-US13682.

15-MAR-1995; 95US-0404228.

23-OCT-1994; 94US-0328224.

25-OCT-1995; 95ZA-0009033.

(IMMU-) IMMUNOLOGIC PHARM CORP.

Devauux B, Franzen H, Geffer M, Hsu D, Pallard X;
 Rothbard J, Samson M, Shi J, Smilek D;

WPI; 1996-230552/23.

Myelin basic derived peptide(s) and analogs - used in the treatment
 of Multiple Sclerosis, psoriasis, Graves Disease, etc.

Claim 1; Fig 2; 91pp; English.

AA95334-R95374 represent peptides derived from myelin basic protein
 (MBP). Immunisation with MBP can be used to induce experimental allergic
 encephalomyelitis (EAE) in susceptible strains of mice. EAE is a CD4+
 T-cell mediated autoimmune disease which results in demyelination of the
 central nervous system, resulting in paralysis and other neurological
 abnormalities. EAE is a commonly used animal model for human multiple
 sclerosis (MS). These sequences can be used in compositions for
 treating MS in a mammal. The composition acts to down regulate the
 autoimmune response, and may be administered in an amount sufficient to
 prevent the onset of symptoms of MS. The compositions may also be used
 to treat advanced stage MS, especially relapsing-remitting MS, chronic
 progressive MS or benign MS. These peptides may also be used in the
 treatment of other diseases involving myelin autoantigens, including
 diabetes, Graves disease, myasthenia gravis, Good Pasture's syndrome,
 psoriasis, thyroiditis, and rheumatoid arthritis. Peptides derived from
 other myelin autoantigens, such as myelin oligodendrocyte protein (MOG),
 proteolipid protein (PLP), and myelin associated glycoprotein (MAG) can
 be used as alternatives to these MBP peptides in these compositions.

Sequence 20 AA:

Query Match 71.4%; Score 35; DB 17; Length 20;
 Best Local Similarity 71.4%; Pred. No. 5;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 ISRFAMG 7
 :|||:|
 1 Isrlswg 7

RESULT 12

AA943950 standard; peptide; 20 AA.

AA943950;

22-APR-1998 (first entry)

Human myelin basic protein peptide MBP-3 (111-130).

Human; myelin oligodendrocyte glycoprotein; MOG; multiple sclerosis;
 autoimmune response; MBP; myelin basic protein; demyelinating.

Homo sapiens.

WO9735879-A1.

02-OCT-1997.

01-MAY-1996; 96MO-US06072.

28-MAR-1996; 96US-0623406.

(IMMU-) IMMUNOLOGIC PHARM CORP.

Devauux B, Garman RD, Rothbard J, Smilek D, Wallner B;

WPI; 1997-489564/45.

Human myelin oligodendrocyte glyco-protein peptide - useful to treat

PT multiple sclerosis in mammal by down-regulating auto-immune response
XX
XX
PS Claim 101; Page 35; 108pp; English.

CC The present sequence represents a novel peptide of human myelin
CC basic protein (MBP). The peptide of human MOG (myelin oligodendrocyte
CC glycoprotein) and MBP can be used to treat multiple sclerosis in a
CC mammal by down-regulating an autoimmune response in the mammal. They
CC can also be used to diagnose and treat other demyelinating autoimmune
CC diseases in humans, or to prepare antibodies for the detection or
CC diagnosis of autoimmune diseases.

XX
SQ Sequence 20 AA;

Query Match 71.4%; Score 35; DB 18; Length 20;
Best Local Similarity 71.4%; Pred. No. 5;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
:||||:
DB 1 Isrlfswg 7

JIT 13

AAW44066
ID AAW44066 standard; peptide; 20 AA.

XX
AC AAW44066;

XX
DT 22-APR-1998 (first entry)

XX
DE Human myelin basic protein peptide MBP-K (101-120).

XX
KW Human; myelin oligodendrocyte glycoprotein; MOG; multiple sclerosis;
KW autoimmune response; MBP; myelin basic protein; demyelinating.

XX
OS Homo sapiens.

XX
PN MO9735879-A1.

XX
PD 02-OCT-1997.

XX
PF 01-MAY-1996; 96WO-US06072.

XX
PR 28-MAR-1996; 96US-0623406.

XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.

XX
PI Devaux B, Garman RD, Rothbard J, Smilek D, Wallner B;

XX
WPI; 1997-489564/45.

PT Human myelin oligodendrocyte glyco-protein peptide - useful to treat
PT multiple sclerosis in mammal by down-regulating auto-immune response
PS Disclosure; Page 35; 108pp; English.

XX
XX The present sequence represents a novel peptide of human myelin
XX basic protein (MBP). The peptide of human MOG (myelin oligodendrocyte
XX glycoprotein) and MBP can be used to treat multiple sclerosis in a
XX mammal by down-regulating an autoimmune response in the mammal. They
XX can also be used to diagnose and treat other demyelinating autoimmune
XX diseases in humans, or to prepare antibodies for the detection or
XX diagnosis of autoimmune diseases.

XX
SQ Sequence 20 AA;

Query Match 71.4%; Score 35; DB 18; Length 20;
Best Local Similarity 71.4%; Pred. No. 5;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
:||||:
DB 11 Isrlfswg 17

RESULT 14
AAW73618
ID AAW73618 standard; peptide; 20 AA.

XX
AC AAW73618;

XX
DT 18-MAR-1999 (first entry)

XX
DE Human myelin basic protein residues 113-132.

XX
KW Myelin basic protein; MBP; human; immunodominant region; inhibitor;
KW immunosuppressant agent; multiple sclerosis; T cell proliferation;
KW suppressor inducer.

XX
OS Homo sapiens.

XX
PN US5858980-A.

XX
PD 12-JAN-1999.

XX
PF 06-JUN-1995; 95US-0468540.

XX
PR 09-APR-1993; 93US-0046354.

XX
PR 30-MAR-1990; 90US-0502559.

XX
PR 28-FEB-1992; 92US-0843752.

XX
PR 09-APR-1992; 92US-0865318.

XX
PR 06-JUN-1995; 95US-0468540.

XX
PA (AUTO-) AUTOIMMUNE INC.

XX
PI Al-Sabbagh A, Hafler DA, Miller A, Weiner HL;

XX
WPI; 1999-119958/10.

XX
DR New peptides from immunodominant regions of human myelin basic
XX protein - useful as immunosuppressants for treatment of multiple
XX sclerosis

XX
XX This sequence represents a fragment of the human myelin basic protein
XX (MBP). The peptides are immunodominant regions of MBP, and are used as
XX the immunosuppressant agents of the invention. The peptides act on the
XX response to MBP (or tissues containing it), specifically for treatment of
XX multiple sclerosis (MS). They either induce anergy in MBP-reactive
XX T cells or actively suppress these cells by inhibiting their
XX proliferation, possibly by inducing suppressor T cells. The peptides may
XX also be used to detect specific T cells. The MBP fragments do not induce
XX 'global' immunosuppression upon prolonged use (increasing the risk of a
XX patient developing certain malignancies), will not down regulate normal
XX immune responses to pathogenic micro-organisms and are less toxic than
XX drugs currently used for the treatment of MS, such as azathioprine
XX (Imuran), and cyclosporin A.

XX
SQ Sequence 20 AA;

Query Match 71.4%; Score 35; DB 20; Length 20;
Best Local Similarity 71.4%; Pred. No. 5;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
:||||:
DB 1 Isrlfswg 7

RESULT 15
AAB12619

D AAB12619 standard; peptide; 20 AA.
 X AAB12619;
 X
 X 10-NOV-2000 (first entry)
 X
 X Human myelin basic protein peptide 113-132.
 X
 X Human; myelin basic protein; MBP; multiple sclerosis; MS; CNS;
 W immunodominant; chronic inflammatory disease; central nervous system;
 M autoimmune disease; neuroprotective; immunosuppressive; CD8+; CD4+.
 X
 X Homo sapiens.
 S
 X US6077509-A.
 X
 X 20-JUN-2000.
 X
 X 06-JUN-1995; 95US-0469648.
 X
 X 09-APR-1993; 93US-0046354.
 R 30-MAR-1990; 90US-0502559.
 R 28-FEB-1992; 92US-0843752.
 R 09-APR-1992; 92US-0865318.
 X
 X (AUTO-) AUTOIMMUNE INC.
 X
 X Hafler DA, Weiner HL;
 X
 X WPI; 2000-450923/39.
 X
 X
 T New pharmaceutical composition comprising peptide fragments of myelin
 T basic protein, useful in suppressing the symptoms of multiple sclerosis
 T
 S Example 1; Column 14; 26pp; English.
 X
 X The present invention describes a pharmaceutical formulation comprising
 C a peptide having a 30-amino acid sequence (AAB12612) (I) or a segment of
 C (I), provided that the segment comprises at least the 19-amino acid
 C sequence given in AAB12613. The pharmaceutical formulation has
 C neuroprotective and immunosuppressive activities. The compositions and
 C peptides from the present invention are useful in suppressing multiple
 C sclerosis, which is due to the induction of CD8+ suppressor T-cells.
 C The peptides are also useful in identifying CD4+ T-cells reactive with
 C myelin basic protein (MBP). The peptides are also useful in identifying
 C individuals with T-cells reactive to MBP. Oral administration of MBP
 C suppressed central nervous system (CNS) inflammation in actively
 C induced experimental allergic encephalomyelitis (EAE). There was a
 C decreased inflammation in both the parenchyma and meninges when cells
 C from MBP-fed animals were transferred. This suppression was observed
 C when CD4+ depleted, but not CD8+. The present sequence represents a
 C human MBP peptide fragment, which is used in an example from the
 C present invention.
 X
 X Sequence 20 AA;
 Q

Query Match 71.4%; Score 35; DB 21; Length 20;
 Best Local Similarity 71.4%; Pred. No. 5;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 ISRRANG 7
 :|||:|
 b 1 ISRFWSG 7

Search completed: October 18, 2001, 16:51:16
 Job time: 41 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2001, 16:51:41 ; Search time 18.9 Seconds
(without alignments)
9.805 Million cell updates/sec

Title: US-09-646-579-2

Perfect score: 49

Sequence: 1 ISRFAMGEV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 segs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Minimum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents_AA: *
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/1aa/PCOMB_COMB.pep.*
6: /cgn2_6/prodata/2/1aa/beckfilles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	71.4	9	3	US-08-159-339A-1078
2	35	71.4	12	2	US-08-735-253-1
3	35	71.4	20	2	US-08-468-540B-11
4	35	71.4	20	3	US-08-297-395-23
5	35	71.4	168	6	5194425-4
6	35	71.4	170	1	US-08-227-372-1
7	35	71.4	170	2	US-08-327-357A-1
8	35	71.4	170	3	US-08-470-397-1
9	35	71.4	170	3	US-09-007-520-1
10	35	71.4	170	3	US-08-462-351-3
11	35	71.4	170	4	US-09-055-263-1
12	35	71.4	170	6	5194425-3
13	35	71.4	170	6	5468481-3
14	35	71.4	171	2	US-08-781-122-2
15	35	71.4	171	4	US-09-137-759-2
16	34	69.4	10	2	US-08-735-253-6
17	34	69.4	10	2	US-08-735-253-7
18	34	69.4	11	2	US-08-735-253-12
19	34	69.4	118	5	PCT-US94-07926-40
20	34	69.4	118	5	PCT-US94-07926-40
21	34	69.4	138	2	US-08-888-497-32
22	34	69.4	138	5	PCT-US94-07926-32
23	34	69.4	138	5	PCT-US94-07926-32
24	33	67.3	1088	2	US-08-633-768A-4
25	33	65.3	239	2	US-08-933-756C-22
26	33	65.3	239	4	US-09-234-613-22
27	32	65.3	872	1	US-08-451-715A-8

28	32	65.3	1621	4	US-08-972-927-3	Sequence 3, Appl
29	30	61.2	111	4	US-08-466-886-36	Sequence 36, Appl
30	30	61.2	111	4	US-08-469-617-36	Sequence 36, Appl
31	30	61.2	263	1	US-08-362-670B-32	Sequence 32, Appl
32	30	61.2	263	5	US-08-333-576C-32	Sequence 32, Appl
33	30	61.2	263	5	PCT-US94-14030A-32	Sequence 32, Appl
34	30	61.2	321	1	US-08-362-670B-26	Sequence 26, Appl
35	30	61.2	321	5	US-08-333-576C-26	Sequence 26, Appl
36	30	61.2	321	5	PCT-US94-14030A-26	Sequence 26, Appl
37	30	61.2	461	5	PCT-US93-12464-2	Sequence 26, Appl
38	30	61.2	578	1	US-08-653-740-3	Sequence 2, Appl
39	30	61.2	578	2	US-09-073-594-3	Sequence 3, Appl
40	30	61.2	578	3	US-09-275-925-3	Sequence 3, Appl
41	30	61.2	623	1	US-08-653-740-7	Sequence 7, Appl
42	30	61.2	623	2	US-09-073-594-7	Sequence 7, Appl
43	30	61.2	623	1	US-09-275-925-7	Sequence 7, Appl
44	30	61.2	636	1	US-08-653-740-5	Sequence 5, Appl
45	30	61.2	636	2	US-09-073-594-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-159-339A-1078
Sequence 1078, Application US/08159339A
Patent No. 6037135
GENERAL INFORMATION:
APPLICANT: Kubo, Ralph T.
APPLICANT: Grey, Howard M.
APPLICANT: Sette, Alessandro
APPLICANT: Celis, Esteban
TITLE OF INVENTION: HLA Binding peptides and Their
NUMBER OF SEQUENCES: 1254
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,339A
FILING DATE: 29-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,666
FILING DATE: 07-AUG-1992
APPLICATION NUMBER: US 08/027,746
FILING DATE: 05-MAR-1993
APPLICATION NUMBER: US 08/103,396
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-00503005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX:
INFORMATION FOR SEQ ID NO: 1078:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-159-339A-1078

Query Match

71.4%; Score 35; DB 3; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.5e+05;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;y 1 ISRFAMG 7
:|||||
db 2 LSRFSWG 8

RESULT 2

US-08-735-253-1
Sequence 1, Application US/08735253
Patent No. 5942491

GENERAL INFORMATION:

APPLICANT: Root-Bernstein, Robert S.

TITLE OF INVENTION: Methods and Compositions for Treating

TITLE OF INVENTION: Arthritis

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dressler, Goldsmith, Milmanow & Katz, Ltd.

STREET: 180 N. Stetson

CITY: Chicago

STATE: Illinois

COUNTRY: U.S.A.

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/735,253

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: No. 5942491thrup, Thomas E.

REGISTRATION NUMBER: 33,268

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5400

TELEFAX: (312) 616-5460

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-735-253-1

Query Match

Best Local Similarity 71.4%; Score 35; DB 2; Length 12;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

y 1 ISRFAMG 7
:|||||
db 2 LSRFSWG 8

RESULT 3

US-08-468-540B-11
Sequence 11, Application US/08468540B
Patent No. 5858980

GENERAL INFORMATION:

APPLICANT: Weiner, Howard

APPLICANT: Hafner, David

APPLICANT: Miller, Ariel

APPLICANT: Al-Sabbagh, Ahmad

TITLE OF INVENTION: SUPPRESSION OF T-CELL PROLIFERATION

TITLE OF INVENTION: USING PEPTIDE FRAGMENTS OF MYELIN BASIC PROTEIN

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby & Darby P.C.

STREET: 805 Third Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,540B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Jacobs, Seth H

REGISTRATION NUMBER: 32,140

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-527-7700

TELEFAX:

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 5858980e

US-08-468-540B-11

Query Match

Best Local Similarity 71.4%; Score 35; DB 2; Length 20;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

y 1 ISRFAMG 7
:|||||
db 1 LSRFSWG 7

RESULT 4

US-08-297-395-23
Sequence 23, Application US/08297395A
Patent No. 6039947

GENERAL INFORMATION:

APPLICANT: Howard L. Weiner

APPLICANT: David A. Hafner

TITLE OF INVENTION: PEPTIDES DERIVED FROM IMMUNODOMINANT

TITLE OF INVENTION: EPITOPES OF MYELIN BASIC PROTEIN

FILE REFERENCE: 1010/057230US

CURRENT APPLICATION NUMBER: US/08/297,395A

CURRENT FILING DATE: 1994-08-11

EARLIER APPLICATION NUMBER: 08/059,189

EARLIER FILING DATE: 1993-05-06

EARLIER APPLICATION NUMBER: 07/502,559

EARLIER FILING DATE: 1990-03-30

EARLIER APPLICATION NUMBER: PCT/US88/02139

EARLIER FILING DATE: 1988-06-24

EARLIER APPLICATION NUMBER: 07/065,734

EARLIER FILING DATE: 1987-06-24

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 23

LENGTH: 20

TYPE: PRT

ORGANISM: Homo sapiens
US-08-297-395-23

Query Match 71.4%; Score 35; DB 3; Length 20;
Best Local Similarity 71.4%; Pred. No. 2.6;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
:||||:
Db. 1 LSFRSMG 7

RESULT 5
5194425-4
Patent No. 5194425
APPLICANT: SHARMA, SOMESH D.; LERCH, L. BERNARD; CLARK,
BRIAN R.
TITLE OF INVENTION: MHC-MEDIATED TOXIC CONJUGATES USEFUL IN
IMPROVING AUTOIMMUNITY
NUMBER OF SEQUENCES: 9
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/367,751
FILING DATE: 21-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 210,594
FILING DATE: 23-JUN-1988
SEQ ID NO: 4
LENGTH: 168
5194425-4

Query Match 71.4%; Score 35; DB 6; Length 168;
Best Local Similarity 71.4%; Pred. No. 2.2;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
:||||:
Db. 109 LSFRSMG 115

RESULT 6
US-08-227-372-1
Sequence 1, Application US/08227372
Patent No. 5763585
GENERAL INFORMATION:
APPLICANT: Nag, Bishwajit
TITLE OF INVENTION: PURIFICATION AND CHARACTERIZATION OF
NUMBER OF SEQUENCES: 4
TITLE OF INVENTION: MHC-PEPTIDE COMPLEXES USEFUL IN AMELIORATING AUTOIMMUNITY
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,372
FILING DATE: 14-APR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/136,216
FILING DATE: 13-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 14058-32-1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..170
OTHER INFORMATION: /note="Myelin basic protein"
US-08-227-372-1

Query Match 71.4%; Score 35; DB 1; Length 170;
Best Local Similarity 71.4%; Pred. No. 2.2;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
:||||:
Db. 111 LSFRSMG 117

RESULT 7
US-08-327-357A-1
Sequence 1, Application US/08327357A
Patent No. 5817629
GENERAL INFORMATION:
APPLICANT: WARREN, Kenneth G.
APPLICANT: CATZ, Ingrid
TITLE OF INVENTION: PEPTIDE SPECIFICITY OF ANTI-MYELIN BASIC
TITLE OF INVENTION: PROTEIN AND THE ADMINISTRATION OF MYELIN BASIC PROTEIN
TITLE OF INVENTION: PEPTIDES TO MULTIPLE SCLEROSIS PATIENTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/327,357A
FILING DATE: 21-OCT-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/798,099
FILING DATE: 27-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CA 2,053,799-0
FILING DATE: 22-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 27052-115469
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-8300
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:

ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: human myelin basic protein
US-08-327-357A-1

Query Match 71.4%; Score 35; DB 2; Length 170;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
:111:11
DB 111 LSFRSMG 117

RESULT 8
US-08-470-397-1
Sequence 1, Application US/08470397
Patent No. 6007820

GENERAL INFORMATION:
APPLICANT: Nag, Bishwajit
TITLE OF INVENTION: Purification and Characterization of
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Stewart Street Tower, 20th
FLOOR
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,397
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/136,216
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,372
FILING DATE: 14-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 14058-32-1-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO

FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - N-acetyl-alanine"
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Ala or absent"

NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Arg or His"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Ser or Gly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 16
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 40
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Leu or Ile"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 46
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Ser or Gly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 59
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Gly or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 66
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Thr or Ala"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 75
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Ala or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 76
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Cln or absent"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 80
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Pro or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 141
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Leu or Phe"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 144
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - His or Val"
FEATURE:
NAME/KEY: Protein
LOCATION: 1..170
OTHER INFORMATION: /note- "Myelin basic protein"

US-08-470-397-1

Query Match 71.4%; Score 35; DB 3; Length 170;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
:111:11
DB 111 LSFRSMG 117

RESULT 9
US-09-007-520-1
Sequence 1, Application US/09007520
Patent No. 6103696
GENERAL INFORMATION:
APPLICANT: WARREN, Kenneth G.
ATTORNEY/AGENT INFORMATION:
NAME: Catz, Ingrid
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 27052-115469
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: human myelin basic protein
US-09-007-520-1

Query Match 71.4%; Score 35; DB 3; Length 170;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISRFAMG 7
DB 111 ISRFAMG 117

RESULT 10
US-08-462-351-3
Sequence 3, Application US/08462351
Patent No. 6106840
GENERAL INFORMATION:
APPLICANT: Sharma, Somesh D.
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Brian R.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 27052-115469
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: human myelin basic protein
US-09-007-520-1

TITLE OF INVENTION: MHC Conjugates Useful in Ameliorating
TITLE OF INVENTION: Autoimmunity
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,351
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/210,594
FILING DATE: 23-JUN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,084
FILING DATE: 30-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/690,840
FILING DATE: 23-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/869,293
FILING DATE: 14-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Parent, Annette S.
REGISTRATION NUMBER: 42,058
REFERENCE/DOCKET NUMBER: 014058-00024205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..170
OTHER INFORMATION: /note="myelin basic protein (MBP)"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /product="OTHER"
OTHER INFORMATION: /note="Xaa - N-acetyl-alanine"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /product="OTHER"
OTHER INFORMATION: /note="Ala at position 3 may be present"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /product="OTHER"
OTHER INFORMATION: /note="Xaa - Arg or His"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /product="OTHER"
OTHER INFORMATION: /note="Ser or Gly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 16

OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 40
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Leu or Ile"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 46
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Ser or Gly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 59
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Gly or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 66
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Thr or Ala"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 75
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Ala or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 76
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Gln at position 76 may be present or absent"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 80
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Pro or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 107
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Arg or Met"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 122
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Lys or Arg"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 141
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - Leu or Phe"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 144
OTHER INFORMATION: /product- "OTHER"
OTHER INFORMATION: /note- "Xaa - His or Val"
US-08-462-351-3

Query Match 71.4%; Score 35; DB 3; Length 170;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
DB 111 LSFSWG 117

RESULT 11
US-09-055-263-1
; Sequence 1, Application US/09055263

Patent No. 6252040
GENERAL INFORMATION:
APPLICANT: WARREN, Kenneth G.
TITLE OF INVENTION: PEPTIDE SPECIFICITY OF ANTI-MYELIN BASIC PROTEIN AND THE ADMINISTRATION OF MYELIN BASIC PROTEIN
TITLE OF INVENTION: PEPTIDES TO MULTIPLE SCLEROSIS PATIENTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/055,263
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/327,357
FILING DATE: 21-OCT-1994
APPLICATION NUMBER: US 07/798,099
FILING DATE: 27-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CA 2,053,799-0
FILING DATE: 22-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Innen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 27052-115469
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: human myelin basic protein
US-09-055-263-1

Query Match 71.4%; Score 35; DB 4; Length 170;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7
DB 111 LSFSWG 117

RESULT 12
5194425-3
Patent No. 5194425
APPLICANT: SHARMA, SOMESH D.; LERCH, L. BERNARD; CLARK, BRIAN R.
TITLE OF INVENTION: MHC-MEDIATED TOXIC CONJUGATES USEFUL IN AMELIORATING AUTOIMMUNITY
NUMBER OF SEQUENCES: 9
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/367,751
FILING DATE: 21-JUN-1989

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 210,594
FILING DATE: 23-JUN-1988
SEQ ID NO: 3
LENGTH: 170
5194425-3

Query Match 71.4%; Score 35; DB 6; Length 170;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
Db 111 LSRSWG 117

RESULT 13
5468481-3

PATENT NO. 5468481
APPLICANT: SHARMA, SOMESH D.; CLARK, BRIAN R.; LERCH, BERNARD L.
TITLE OF INVENTION: MHC CLASS II-PEPTIDE CONJUGATES USEFUL
IN AMELIORATING AUTOIMMUNITY

NUMBER OF SEQUENCES: 7
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,293
FILING DATE: 14-APR-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 690,840
FILING DATE: 23-APR-1991
APPLICATION NUMBER: 576,084
FILING DATE: 30-AUG-1990
APPLICATION NUMBER: 210,594
FILING DATE: 23-JUN-1988
APPLICATION NUMBER: 635,840
FILING DATE: 28-DEC-1998
APPLICATION NUMBER: 367,751
FILING DATE: 21-JUN-1989
SEQ ID NO: 3
LENGTH: 170
5468481-3

Query Match 71.4%; Score 35; DB 6; Length 170;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
Db 111 LSRSWG 117

RESULT 14
US-08-781-122-2

Sequence 2, Application US/08781122
Patent No. 5948764
GENERAL INFORMATION:

APPLICANT: Gaur, Amitabh
APPLICANT: Conlon, Paul J.
TITLE OF INVENTION: METHODS FOR TREATMENT OF MULTIPLE SCLEROSIS
TITLE OF INVENTION: UTILIZING PEPTIDE ANALOGUES OF HUMAN MYELIN BASIC PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,122
FILING DATE: 09-JAN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Makl, David J.
REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 690068.418C1
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 171 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-781-122-2

Query Match 71.4%; Score 35; DB 2; Length 171;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
Db 112 LSRSWG 118

RESULT 15
US-09-137-759-2

Sequence 2, Application US/09137759
Patent No. 6251396
GENERAL INFORMATION:

APPLICANT: Gaur, Amitabh
APPLICANT: Conlon, Paul J.

APPLICANT: Ling, Nicholas C.
APPLICANT: Staehlin, Theophil

APPLICANT: Crowe, Paul D.
TITLE OF INVENTION: METHODS FOR TREATMENT OF MULTIPLE SCLEROSIS USING

FILE REFERENCE: 690068.405C1
CURRENT APPLICATION NUMBER: US/09/137,759
CURRENT FILING DATE: 1998-08-20

NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2
LENGTH: 171
TYPE: PRT

ORGANISM: Homo sapiens
US-09-137-759-2

Query Match 71.4%; Score 35; DB 4; Length 171;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
Db 112 LSRSWG 118

Search completed: October 18, 2001, 16:51:42
Job time: 67 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2001, 16:52:11 ; Search time 22.89 Seconds
(without alignments)
29.951 Million cell updates/sec

Title: US-09-646-579-2

Perfect score: 49

Sequence: 1 ISRFAMGEV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

lmm DB seq length: 0
lmm DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	134	2	B35119
2	45	91.8	373	2	T47115
3	45	91.8	375	2	T35015
4	40	81.6	250	2	E82821
5	40	81.6	449	2	T29425
6	38	77.6	126	2	H75300
7	37	75.5	289	2	S77232
8	37	75.5	394	2	I51681
9	36	73.5	230	1	MNTV16
10	36	73.5	230	1	MNTV16
11	36	73.5	230	1	MNTV16
12	36	73.5	230	1	MNTV16
13	36	73.5	230	1	MNTV16
14	36	73.5	230	1	MNTV16
15	36	73.5	230	1	MNTV16
16	36	73.5	230	1	MNTV16
17	36	73.5	230	1	MNTV16
18	36	73.5	230	1	MNTV16
19	36	73.5	230	1	MNTV16
20	36	73.5	230	1	MNTV16
21	36	73.5	230	1	MNTV16
22	36	73.5	230	1	MNTV16
23	36	73.5	230	1	MNTV16
24	36	73.5	230	1	MNTV16
25	36	73.5	230	1	MNTV16
26	36	73.5	230	1	MNTV16
27	36	73.5	230	1	MNTV16
28	36	73.5	230	1	MNTV16
29	36	73.5	230	1	MNTV16

30	35	71.4	639	2	JC4881	polyvinyl-alcohol
31	34	69.4	138	2	A49959	phospholipase A2 (
32	34	69.4	237	1	MNTV16	nonstructural prot
33	34	69.4	237	1	MNTV16	nonstructural prot
34	34	69.4	488	2	G81213	conserved hypotet
35	34	69.4	490	2	C81790	conserved hypotet
36	34	69.4	592	2	E57032	carbon starvation
37	34	69.4	611	2	C58893	NADH dehydrogenase
38	33	67.3	160	2	T22205	hypothetical prote
39	33	67.3	254	2	E69533	hypothetical prote
40	33	67.3	254	2	F69547	hypothetical prote
41	33	67.3	328	2	T15469	nucleotide-binding
42	33	67.3	336	2	T41118	hypothetical prote
43	33	67.3	449	2	B69759	sodium/proline sym
44	33	67.3	454	2	D84931	chromosomal replic
45	33	67.3	463	2	E86464	hypothetical prote

ALIGNMENTS

RESULT 1
B35119
4-carboxymuconolactone decarboxylase (EC 4.1.1.44) - Acinetobacter calcoaceticus
C:Species: Acinetobacter calcoaceticus 27-Jul-1990 #sequence_revision 08-Sep-1997
C:Date: 27-Jul-1990 #sequence_revision 08-Sep-1997
C:Accession: B35119
R:Hartrick, C.; Neidle, E.L.; Ngai, K.L.; Ornstien, L.N.
J. Bacteriol. 172, 956-966, 1990
A:Title: DNA sequences of genes encoding Acinetobacter calcoaceticus protocatechuate
Y divergence.
A:Reference number: A35119, MID:90130333
A:Accession: B35119
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-134 <HAR>
A:Cross-References: GB:M3798; NID:q141771; PID:q141772
C:Superfamily: 4-carboxymuconolactone decarboxylase
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 100.0%; Score 49; DB 2; Length 134;

Best Local Similarity 100.0%; Pred. No. 0.065; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMGEV 9
DB 41 ISRFAMGEV 49

RESULT 2
T47115
probable 4-carboxymuconolactone decarboxylase / 3-oxoadipate enol-lactone hydrolase [C:Species: Streptomyces sp.
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: T47115
R:Yang, K.; Iwagami, S.; Davies, J.E.
submitted to the EMBL Data Library, May 1999
A:Description: A protocatechuate catabolic gene cluster cloned from Streptomyces sp.
A:Reference number: Z24354
A:Accession: T47115
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-373 <YAN>
A:Cross-References: EMBL:AF109386; PIDN:AND40815.1
C:Experimental source: strain 2065
C:Genetics:
A:Gene: pcal

Query Match 91.8%; Score 45; DB 2; Length 373;
Best Local Similarity 77.8%; Pred. No. 0.93; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMGEV 9
 Db 288 ISRYAMGEI 296

RESULT 3

Probable 3-oxoadipate enol-lactone hydrolase / 4-carboxymuconolactone decarboxylase - St
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 03-Dec-1999
 C:Accession: T35015
 R:Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, June 1999
 A:Reference number: 221565
 A:Accession: T35015
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-375 <SEE>
 A:Cross-references: EMBL:AL079355; PIDN:CA845571.1; GSPDB:GN00070; SCQEDB:SCAC6.07C
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: pcal; SCQEDB:SCAC6.07C

Query Match 91.8%; Score 45; DB 2: Length 375;
 Best Local Similarity 77.8%; Pred. No. 0.93;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMGEV 9
 Db 286 ISRYAMGEI 294

RESULT 4

NADH-ubiquinone oxidoreductase, NQO5 subunit XF0307 [Imported] - Xylella fastidiosa (str
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: E82821
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MIDID:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: E82821
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-250 <SIM>
 A:Cross-references: GB:AE003884; GB:AE003849; NID:99105127; PIDN:AAF3118.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 Submitted to Genbank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0307

Query Match 81.6%; Score 40; DB 2: Length 250;
 Best Local Similarity 75.0%; Pred. No. 5;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ISRFAMGEV 8
 Db 90 VGRFAMGEV 97

RESULT 5

4-carboxymuconolactone decarboxylase homolog - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C:Accession: T29425
 R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, July 1998
 A:Reference number: 220619
 A:Accession: T29425
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-449 <PAR>
 A:Cross-references: EMBL:AL031155; NID:el313489; PID:el313496; PIDN:CAA20070.1
 C:Genetics:
 A:Note: SC3A7.07

Query Match 81.6%; Score 40; DB 2: Length 449;
 Best Local Similarity 55.6%; Pred. No. 8.9;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMGEV 9
 Db 361 LTRYAMGEI 369

RESULT 6

Probable laccoglutarathione lyase - Deinococcus radiodurans (strain RI)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: H75300
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RI.
 A:Reference number: A75250; MIDID:20036896
 A:Accession: H75300
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-126 <WHI>
 A:Cross-references: GB:AE002054; GB:AE000513; NID:96460010; PIDN:AAF11756.1; PID:9646
 A:Experimental source: strain RI
 C:Genetics:
 A:Gene: DR2208
 A:Map position: 1

Query Match 77.6%; Score 38; DB 2: Length 126;
 Best Local Similarity 62.5%; Pred. No. 5.8;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMGEV 8
 Db 94 VRRFAMGEV 101

RESULT 7

hypothetical protein sl11348 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S77232
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
 O.; K., Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* 8.

A:Reference number: S74322; MUID:97061201

A:Accession: S77232

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-289 <KAN>

A:Cross-references: EMBL:D90907; GB:AB001339; NID:g1652618; PIDN:BA17566.1; PID:d101829

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Start codon: GTG

Query Match 75.5%; Score 37; DB 2; Length 289;
 Best Local Similarity 85.7%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISRFAMG 7

DB 73 ISRYAMG 79

JLT 8

poly(A) polymerase - African clawed frog

C:Species: *Xenopus laevis* (African clawed frog)

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: I51681

R:Gebauser, F.; Richter, J.D.

Mol. Cell. Biol. 15, 1422-1430, 1995

A>Title: Cloning and characterization of a *Xenopus poly(A)* polymerase.

A:Reference number: I51681; MUID:95166227

A:Accession: I51681

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-394 <GB>

A:Cross-references: EMBL:U23456; NID:g755871; PIDN:AAA64708.1; PID:g755872

C:Genetics:

Query Match 75.5%; Score 37; DB 2; Length 394;
 Best Local Similarity 55.6%; Pred. No. 27;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISRFAMGEV 9

DB 376 LAMFAMGEI 384

UJT 9

nonstructural protein NS1 - Influenza A virus (strain A/duck/Alberta/60/76)

C:Species: Influenza A virus

C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 16-Jul-1999

C:Accession: A04092

R:Baez, M.; Zazra, J.J.; Elliott, R.M.; Young, J.F.; Palese, P.

Virology 113, 397-402, 1981

A>Title: Nucleotide sequence of the Influenza A/duck/Alberta/60/76 virus NS RNA: conserv

A:Reference number: A04092; MUID:81276929

A:Accession: A04092

A:Molecule type: genomic RNA

A:Residues: 1-230 <BAE>

A:Cross-references: GB:J02105; GB:M17071; NID:g324783; PIDN:AAA43509.1; PID:g324785

C:Genetics:

A:Map position: segment 8
 C:Superfamily: Influenza virus nonstructural protein NS1
 C:Keywords: alternative splicing

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISRFAMG 7

DB 198 IORFAMG 204

RESULT 10

nonstructural protein NS1 - Influenza A virus (strain A/mallard/Alberta/88/76)

C:Species: Influenza A virus

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 08-Apr-1994

C:Accession: A32663

R:Treanor, J.J.; Snyder, M.H.; London, W.T.; Murphy, B.R.

Virology 171, 1-9, 1989

A>Title: The B allele of the NS gene of avian influenza viruses, but not the A allele

A:Reference number: A32662; MUID:89299445

A:Accession: A32663

A:Molecule type: genomic RNA

A:Residues: 1-230 <TRE>

C:Genetics:

A:Gene: NS1

A:Map position: segment 8

C:Superfamily: Influenza virus nonstructural protein NS1

C:Keywords: alternative splicing; nonstructural protein

Query Match 73.5%; Score 36; DB 1; Length 230;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISRFAMG 7

DB 198 IORFAMG 204

RESULT 11

nonstructural protein NS1 - Influenza A virus (strain A/mallard/Alberta/82/78)

C:Species: Influenza A virus

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 08-Apr-1994

C:Accession: C32663

R:Treanor, J.J.; Snyder, M.H.; London, W.T.; Murphy, B.R.

Virology 171, 1-9, 1989

A>Title: The B allele of the NS gene of avian influenza viruses, but not the A allele

A:Reference number: A32662; MUID:89299445

A:Accession: C32663

A:Molecule type: genomic RNA

A:Residues: 1-230 <TRE>

C:Genetics:

A:Gene: NS1

A:Map position: segment 8

C:Superfamily: Influenza virus nonstructural protein NS1

C:Keywords: alternative splicing; nonstructural protein

Query Match 73.5%; Score 36; DB 1; Length 230;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISRFAMG 7

DB 198 IORFAMG 204

RESULT 12

nonstructural protein NS1 - Influenza A virus (strain A/pintail/Alberta/121/79)

C:Species: Influenza A virus

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 08-Apr-1994

C:Accession: E32663

R:Treanor, J.J.; Snyder, M.H.; London, W.T.; Murphy, B.R.

Virology 171, 1-9, 1989

A:Title: The B allele of the NS gene of avian Influenza viruses, but not the A allele, A
A:Reference number: A32662; MUID:89299445

A:Accession: E32663

A:Molecule type: genomic RNA

A:Residues: 1-230 <TR>

C:Genetics:

A:Gene: NS1

A:Map position: segment 8

C:Superfamily: Influenza virus nonstructural protein NS1

C:Keywords: alternative splicing; nonstructural protein

Query Match

Best Local Similarity 73.5%; Score 36; DB 1; Length 230;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
1 11111
198 IORFAMG 204

RESULT 13

MRIVAB

nonstructural protein NS1 - Influenza A virus (strain A/Pintail/Albertain/358/79)

C:Species: Influenza A virus

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 08-Apr-1994

C:Accession: G32663

R:Renov, J.J.; Snyder, M.H.; London, W.T.; Murphy, B.R.

Virolgy 171, 1-9, 1989

A:Title: The B allele of the NS gene of avian Influenza viruses, but not the A allele, A

A:Reference number: A32662; MUID:89299445

A:Accession: G32663

A:Molecule type: genomic RNA

A:Residues: 1-230 <TR>

C:Genetics:

A:Gene: NS1

A:Map position: segment 8

C:Superfamily: Influenza virus nonstructural protein NS1

C:Keywords: alternative splicing; nonstructural protein

Query Match

Best Local Similarity 73.5%; Score 36; DB 1; Length 230;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
1 11111
198 IORFAMG 204

RESULT 14

AA5575

nonstructural protein NS1 - Influenza A virus (strain A/Turkey/Miscosin/68 [H5N9])

C:Species: Influenza A virus

C:Date: 22-Apr-1993 #sequence_revision 22-Jul-1994 #text_change 19-May-2000

C:Accession: AA5575

R:Perdue, M.L.

Virus Res. 23, 223-240, 1992

A:Title: Naturally occurring NS gene variants in an avian Influenza virus isolate.

A:Reference number: AA5575; MUID:92327829

A:Accession: AA5575

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-230 <PER>

A:Cross-references: GB:S39287

A>Note: the authors translated the codon CTG for residue 141 as Gly, CTG for residue 147

C:Superfamily: Influenza virus nonstructural protein NS1

Query Match

Best Local Similarity 73.5%; Score 36; DB 2; Length 230;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
1 11111
Db 198 IORFAMG 204

RESULT 15

T30305

dnaA protein - Lactococcus lactis (fragment)

C:Species: Lactococcus lactis

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 03-Nov-2000

C:Accession: T30305

R:El-Karoui, M.; Ehrlich, D.; Gruss, A.

Proc. Natl. Acad. Sci. U.S.A. 95, 626-631, 1998

A:Title: Identification of the lactococcal exonuclease/recombinase and its modulation

A:Reference number: 220817; MUID:98118563

A:Accession: T30305

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-272 <FLK>

C:Cross-references: EMBL:U76424; MID:92909712; PID:92909713; PIDN:ANCI2963.1

C:Genetics:

A:Gene: dnaA

C:Superfamily: replication initiation protein dnaA

Query Match

Best Local Similarity 73.5%; Score 36; DB 2; Length 272;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ISRFAMG 7
1 11111
Db 86 VSRFSWG 92

Search completed: October 18, 2001, 16:52:11
Job time: 96 sec

Thu Oct 18 16:57:44 2001

us-09-646-579-2.rpx
